Part IA — Numbers and Sets

Based on lectures by A. G. Thomason

Notes taken by Dexter Chua

Michaelmas 2014

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

Introduction to number systems and logic
Overview of the natural numbers, integers, real numbers, rational and irrational numbers, algebraic and transcendental numbers. Brief discussion of complex numbers; statement of the Fundamental Theorem of Algebra.

Ideas of axiomatic systems and proof within mathematics; the need for proof; the role of counter-examples in mathematics. Elementary logic; implication and negation; examples of negation of compound statements. Proof by contradiction. [2]

Sets, relations and functions
Union, intersection and equality of sets. Indicator (characteristic) functions; their use in establishing set identities. Functions; injections, surjections and bijections. Relations, and equivalence relations. Counting the combinations or permutations of a set. The Inclusion-Exclusion Principle. [4]

The integers
The natural numbers: mathematical induction and the well-ordering principle. Examples, including the Binomial Theorem. [2]

Elementary number theory
Prime numbers: existence and uniqueness of prime factorisation into primes; highest common factors and least common multiples. Euclid's proof of the infinity of primes. Euclid's algorithm. Solution in integers of $ax + by = c$.

Modular arithmetic (congruences). Units modulo n. Chinese Remainder Theorem. Wilson's Theorem; the Fermat-Euler Theorem. Public key cryptography and the RSA algorithm. [8]

The real numbers
Least upper bounds; simple examples. Least upper bound axiom. Sequences and series; convergence of bounded monotonic sequences. Irrationality of $\sqrt{2}$ and e. Decimal expansions. Construction of a transcendental number. [4]

Countability and uncountability
Definitions of finite, infinite, countable and uncountable sets. A countable union of countable sets is countable. Uncountability of R. Non-existence of a bijection from a set to its power set. Indirect proof of existence of transcendental numbers. [4]

Contents

0 Introduction **3**

1 Proofs and logic **4**
1.1 Proofs 4
1.2 Examples of proofs 5
1.3 Logic 6

2 Sets, functions and relations **7**
2.1 Sets 7
2.2 Functions 8
2.3 Relations 10

3 Division **12**
3.1 Euclid's Algorithm 12
3.2 Primes 15

4 Counting and integers **18**
4.1 Basic counting 18
4.2 Combinations 20
4.3 Well-ordering and induction 22

5 Modular arithmetic **27**
5.1 Modular arithmetic 27
5.2 Multiple moduli 29
5.3 Prime moduli 30
5.4 Public-key (asymmetric) cryptography 32

6 Real numbers **34**
6.1 Construction of numbers 34
6.2 Sequences 39
6.3 Series 43
6.4 Irrational numbers 44
6.5 Euler's number 45
6.6 Algebraic numbers 45

7 Countability **47**

0 Introduction

According to the Faculty, this course is not aimed at teaching you any new knowledge. In particular, the Faculty says

> This course is concerned not so much with teaching you new parts of mathematics ...

Instead, this course is intended to teach you about how to do maths properly. The objective of this course is to start from a few *axioms*, which are assumptions about our system, and try to prove everything rigorously from the axioms.

This is different from how mathematics is usually done in secondary school. In secondary school, we just accept that certain statements are true. For example, probably no one rigorously proved to you that each natural number has a unique prime factorization. In this course, *nothing* is handwaved. Everything will be obtained as a logical consequence of the axioms and previous (proven) results.

In the course, you shouldn't focus too much on the content itself. Instead, the major takeaway should be how we do mathematics. In particular, how we construct and present arguments.

The actual content of this course is rather diverse. As the name suggests, the course touches on numbers and sets. The "numbers" part is mostly basic number theory, which you may have come across if you have participated in mathematical olympiads. We also study some properties of real numbers, but the actual serious study of real numbers is done in IA Analysis I.

The "sets" part starts with some basic definitions of sets, functions and relations, which are really important and will crop up in all courses you will encounter. In some sense, these form the language in which mathematics is written. At the end of the course, we will touch on countability, which tells us how "big" a set is.

1 Proofs and logic

1.1 Proofs

As the first course in pure mathematics, we will have an (informal) look at proofs and logic.

In mathematics, we often want to *prove* things. This is different from most other disciplines. For example, in science, we perform experiments to convince ourselves that our theory is correct. However, no matter how many experiments we do, we still cannot be absolutely sure that our theory is correct. For example, Newtonian mechanics was believed to be correct for a long time, but is nowadays only considered to be an approximation of reality.

On the other hand, when we prove a theorem in mathematics, we are completely sure that the theorem is true. Also, when we actually prove a theorem, (hopefully) we can also understand *why* it is true, and gain further insight.

Definition (Proof)**.** A *proof* is a sequence of true statements, without logical gaps, that is a logical argument establishing some conclusion.

To prove things, we need to start from some assumptions. These assumptions are known as *axioms*. When we call something an axiom, it does *not* mean that we take these statements to be true without questioning. Instead, we are saying "if we assume these axioms, then these results hold". Two people can disagree on what the axioms are and still be friends.

We also tend to define concepts as unambiguously as possible. Of course, just like a dictionary cannot define all words without being circular, we do not define *everything* in mathematics. To prove things, we have to start somewhere, with some agreed assumptions (*axioms*). We also don't define everything rigorously (or else how could one start speaking?).

In mathematics, we are often concerned about truth. Often, we only care about statements that can take some truth value.

Definition (Statement)**.** A *statement* is a sentence that can have a true value.

Example. The following are statements:

(i) There are infinitely many primes of the form $n^2 + 1$

(ii) There is always a prime number between n and $2n$

(iii) There is no computer program that can factorize an n-digit number in n^3 steps

(iv) For every polynomial $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$, where a_i are complex numbers, $n \geq 1$, $a_n \neq 0$, there exists (possibly complex) z such that $p(z) = 0$

(v) $m \times n = n \times m$ for all natural numbers n, m

(vi) $2 + 2 = 4$

The current status (as of 2015) of these statements are:

(i) No one has a proof but it is probably true

(ii) This is known to be true

(iii) No one knows (related to P = NP problem)

(iv) This is known to be true (Fundamental Theorem of Algebra)

(v) This is known to be true

(vi) This is known to be true (obviously — does this need to be proved?)

1.2 Examples of proofs

Apart from having a proof, it is very important that a proof is *correct*. Here we will look at some examples of proofs and non-proofs.

We first start with a simple example.

Proposition. For all natural numbers n, $n^3 - n$ is a multiple of 3.

Proof. We have $n^3 - n = (n-1)n(n+1)$. One of the three consecutive integers is divisible by 3. Hence so is their product. □

Proposition. If n^2 is even, then so is n.

Proof. If n is even, then $n = 2k$ for some integer k. Then $n^2 = 4k^2$, which is even. □

This is incorrect! We wanted to proof "n^2 is even" $\Rightarrow$ "n is even", but what we proved is "n is even" $\Rightarrow$ "n^2 is even", which are distinct statements.

Instead, a correct proof is as follows:

Proof. Suppose n is odd. Then $n = 2k + 1$ for some integer k. Then $n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2) + 1$, which is odd. This contradicts our assumption that n^2 is even. □

This is an example of *proof by contradiction*. We assume what we want to prove is false, and show that this leads to nonsense.

Proposition. The solutions to $x^2 - 5x + 6 = 0$ are $x = 2$ and $x = 3$.

Note that these are actually 2 different statements:

(i) $x = 2$ and $x = 3$ are solutions

(ii) There are no other solutions

We can write this as an "if and only if" statement: x is a solution if and only if $x = 2$ or $x = 3$. Alternatively, we say "x is a solution iff $x = 2$ or $x = 3$"; or "x is a solution $\Leftrightarrow$ $x = 2$ or $x = 3$".

Proof.

(i) If $x = 2$ or $x = 3$, then $x - 2 = 0$ or $x - 3 = 0$. So $(x - 2)(x - 3) = 0$.

(ii) If $x^2 - 5x + 6 = 0$, then $(x - 2)(x - 3) = 0$. So $x - 2 = 0$ or $x - 3 = 0$. Then $x = 2$ or $x = 3$.

Note that the second direction is simply the first argument reversed. We can write this all in one go:

$$\begin{aligned} x = 3 \text{ or } x = 2 &\Leftrightarrow x - 3 = 0 \text{ or } x - 2 = 0 \\ &\Leftrightarrow (x - 3)(x - 2) = 0 \\ &\Leftrightarrow x^2 - 5x - 6 = 0 \end{aligned}$$

Note that we used the "if and only if" sign between all lines. □

We'll do another non-proof.

Proposition. Every positive number is ≥ 1.

Proof. Let r be the smallest positive real. Then either $r < 1$, $r = 1$ or $r > 1$.

If $r < 1$, then $0 < r^2 < r$. Contradiction. If $r > 1$, then $0 < \sqrt{r} < r$. Contradiction. So $r = 1$. □

Now this is obviously false, since $0.5 < 1$. The problem with this proof is that the smallest positive real need not exist. So we have to make sure we justify all our claims.

1.3 Logic

Mathematics is full of logical statements, which are made of statements and logical connectives. Usually, we use shorthands for the logical connectives.

Let P and Q be statements. Then $P \wedge Q$ stands for "P and Q"; $P \vee Q$ stands for "P or Q"; $P \Rightarrow Q$ stands for "P implies Q"; $P \Leftrightarrow Q$ stands for "P iff Q"; $\neg P$ stands for "not P". The truth of these statements depends on the truth of P and Q . It can be shown by a truth table:

P	Q	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$	$\neg P$
T	T	T	T	T	T	F
T	F	F	T	F	F	F
F	T	F	T	T	F	T
F	F	F	F	T	T	T

Certain logical propositions are equivalent, which we denote using the $\Leftrightarrow$ sign. For example,

$$\neg(P \wedge Q) \Leftrightarrow (\neg P \vee \neg Q),$$

or

$$(P \Rightarrow Q) \Leftrightarrow (\neg P \vee Q) \Leftrightarrow (\neg Q \Rightarrow \neg P).$$

By convention, negation has the highest precedence when bracketing. For example, $\neg P \vee \neg Q$ should be bracketed as $(\neg P) \vee (\neg Q)$.

We also have quantifiers. $(\forall x)P(x)$ means "for all x, $P(x)$ is true", while $(\exists x)P(x)$ means "there exists x such that $P(x)$ is true".

The quantifiers are usually *bounded*, i.e. we write $\forall x \in X$ or $\exists x \in X$ to mean "for all x in the set X" and "there exists x in the set X" respectively.

Quantifiers are negated as follows:

$$\neg(\forall x)P(x) \Leftrightarrow (\exists x)(\neg P(x));$$

$$\neg(\exists x)P(x) \Leftrightarrow (\forall x)(\neg P(x)).$$

2 Sets, functions and relations

In this chapter, we will look at the basic building blocks of mathematics, namely sets, functions and relations. Most of mathematics can be expressed in terms of these notions, and it is helpful to be familiar with the relevant terms.

2.1 Sets

Definition (Set)**.** A *set* is a collection of stuff, without regards to order. Elements in a set are only counted once. For example, if $a = 2, b = c = 1$, then $A = \{a, b, c\}$ has only two members. We write $x \in X$ if x is a member of the set X.

Example. Common sets and the symbols used to denote them:

- $\mathbb{N} = \{1, 2, 3, \cdots\}$ is the natural numbers
- $\mathbb{N}_0 = \{0, 1, 2, \cdots\}$ is the natural numbers with 0
- $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ is the integers
- $\mathbb{Q} = \{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$ is the rational numbers
- $\mathbb{R}$ is the real numbers
- $\mathbb{C}$ is the complex numbers

It is still debated whether 0 is a natural number. Those who believe that 0 is a natural number usually write $\mathbb{N}$ for $\{0, 1, 2, \cdots\}$, and $\mathbb{N}^+$ for the positive natural numbers. However, most of the time, it doesn't matter, and when it does, you should specify it explicitly.

Definition (Equality of sets)**.** A is equal to B, written as $A = B$, if

$$(\forall x)\, x \in A \Leftrightarrow x \in B,$$

i.e. two sets are equal if they have the same elements.

Definition (Subsets)**.** A is a *subset* of B, written as $A \subseteq B$ or $A \subset B$, if all elements in A are in B. i.e.

$$(\forall x)\, x \in A \Rightarrow x \in B.$$

Theorem. $(A = B) \Leftrightarrow (A \subseteq B \text{ and } B \subseteq A)$

Suppose X is a set and P is the property of some elements in x, we can write a set $\{x \in X : P(x)\}$ for the subset of x comprising of the elements for which $P(x)$ is true. e.g. $\{n \in \mathbb{N} : n \text{ is prime}\}$ is the set of all primes.

Definition (Intersection, union, set difference, symmetric difference and power set)**.** Given two sets A and B, we define the following:

- Intersection: $A \cap B = \{x : x \in A \text{ and } x \in B\}$
- Union: $A \cup B = \{x : x \in A \text{ or } x \in B\}$

- Set difference: $A \setminus B = \{x \in A : x \notin B\}$
- Symmetric difference: $A \Delta B = \{x : x \in A \text{ xor } x \in B\}$, i.e. the elements in exactly one of the two sets
- Power set: $\mathcal{P}(A) = \{X : X \subseteq A\}$, i.e. the set of all subsets

New sets can only be created via the above operations on old sets (plus replacement, which says that you can replace an element of a set with another element). One cannot arbitrarily create sets such as $X = \{x : x \text{ is a set and } x \notin x\}$. Otherwise paradoxes will arise.

We have several rules regarding how these set operations behave, which should be intuitively obvious.

Proposition.

- $(A \cap B) \cap C = A \cap (B \cap C)$
- $(A \cup B) \cup C = A \cup (B \cup C)$
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Notation. If A_α are sets for all $\alpha \in I$, then

$$\bigcap_{\alpha \in I} A_\alpha = \{x : (\forall \alpha \in I) x \in A_\alpha\}$$

and

$$\bigcup_{\alpha \in I} A_\alpha = \{x : (\exists \alpha \in I) x \in A_\alpha\}.$$

Definition (Ordered pair)**.** An *ordered pair* (a, b) is a pair of two items in which order matters. Formally, it is defined as $\{\{a\}, \{a, b\}\}$. We have $(a, b) = (a', b')$ iff $a = a'$ and $b = b'$.

Definition (Cartesian product)**.** Given two sets A, B, the *Cartesian product* of A and B is $A \times B = \{(a, b) : a \in A, b \in B\}$. This can be extended to n products, e.g. $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) : x, y, z \in \mathbb{R}\}$ (which is officially $\{(x, (y, z)) : x, y, z \in \mathbb{R}\}$).

2.2 Functions

Definition (Function/map)**.** A *function* (or *map*) $f : A \to B$ is a "rule" that assigns, for each $a \in A$, precisely one element $f(a) \in B$. We can write $a \mapsto f(a)$. A and B are called the *domain* and *co-domain* respectively.

If we wish to be very formal, we can define a function to be a subset $f \subseteq A \times B$ such that for any $a \in A$, there exists a unique $b \in B$ such that $(a, b) \in f$. We then think of $(a, b) \in f$ as saying $f(a) = b$. However, while this might act as a formal definition of a function, it is a terrible way to think about functions.

Example. $x^2 : \mathbb{R} \to \mathbb{R}$ is a function that sends x to x^2. $\frac{1}{x} : \mathbb{R} \to \mathbb{R}$ is not a function since $f(0)$ is not defined. $\pm x : \mathbb{R} \to \mathbb{R}$ is also not a function since it is multi-valued.

It is often helpful to categorize functions into different categories.

Definition (Injective function)**.** A function $f : X \to Y$ is *injective* if it hits everything at most once, i.e.

$$(\forall x, y \in X)\, f(x) = f(y) \Rightarrow x = y.$$

Definition (Surjective function)**.** A function $f : X \to Y$ is *surjective* if it hits everything at least once, i.e.

$$(\forall y \in Y)(\exists x \in X)\, f(x) = y$$

Example. $f : \mathbb{R} \to \mathbb{R}^+ \cup \{0\}$ with $x \mapsto x^2$ is surjective but not injective.

Definition (Bijective function)**.** A function is *bijective* if it is both injective and surjective. i.e. it hits everything exactly once.

Definition (Permutation)**.** A *permutation* of A is a bijection $A \to A$.

Definition (Composition of functions)**.** The *composition* of two functions is a function you get by applying one after another. In particular, if $f : X \to Y$ and $G : Y \to Z$, then $g \circ f : X \to Z$ is defined by $g \circ f(x) = g(f(x))$. Note that function composition is associative.

Definition (Image of function)**.** If $f : A \to B$ and $U \subseteq A$, then $f(U) = \{f(u) : u \in U\}$. $f(A)$ is the *image* of A.

By definition, f is surjective iff $f(A) = B$.

Definition (Pre-image of function)**.** If $f : A \to B$ and $V \subseteq B$, then $f^{-1}(V) = \{a \in A : f(a) \in V\}$.

This is the *pre-image* of the function f, and acts on *subsets* of B. This is defined for any function f. It is important to note that we use the same symbol f^{-1} to denote the *inverse function*, which we will define later, but they are very distinct entities. For example, we will see that the inverse function exists only for bijective functions.

To define the inverse function, we will first need some preliminary definitions.

Definition (Identity map)**.** The *identity map* $\mathrm{id}_A : A \to A$ is defined as the map $a \mapsto a$.

Definition (Left inverse of function)**.** Given $f : A \to B$, a *left inverse* of f is a function $g : B \to A$ such that $g \circ f = \mathrm{id}_A$.

Definition (Right inverse of function)**.** Given $f : A \to B$, a *right inverse* of f is a function $g : B \to A$ such that $f \circ g = \mathrm{id}_B$.

Theorem. The left inverse of f exists iff f is injective.

Proof. ($\Rightarrow$) If the left inverse g exists, then $\forall a, a' \in A, f(a) = f(a') \Rightarrow g(f(a)) = g(f(a')) \Rightarrow a = a'$. Therefore f is injective.

($\Leftarrow$) If f is injective, we can construct a g defined as

$$g : \begin{cases} g(b) = a & \text{if } b \in f(A), \text{ where } f(a) = b \\ g(b) = \text{anything} & \text{otherwise} \end{cases}.$$

Then g is a left inverse of f. □

Theorem. The right inverse of f exists iff f is surjective.

Proof. ($\Rightarrow$) We have $f(g(B)) = B$ since $f \circ g$ is the identity function. Thus f must be surjective since its image is B.

($\Leftarrow$) If f is surjective, we can construct a g such that for each $b \in B$, pick one $a \in A$ with $f(a) = b$, and put $g(b) = a$. □

(Note that to prove the second part, for each b, we need to *pick* an a such that $f(a) = b$. If B is infinite, doing so involves making infinite arbitrary choices. Are we allowed to do so?

To make infinite choices, we need to use the *Axiom of choice*, which explicitly says that this is allowed. In particular, it says that given a family of sets A_i for $i \in I$, there exists a *choice function* $f : I \to \bigcup A_i$ such that $f(i) \in A_i$ for all i.

So can we prove the theorem without the Axiom of Choice? The answer is no. This is since if we assume surjective functions have inverses, then we can prove the Axiom of Choice.

Assume any surjective function f has a right inverse. Given a family of non-empty sets A_i for $i \in I$ (wlog assume they are disjoint), define a function $f : \bigcup A_i \to I$ that sends each element to the set that contains the element. This is surjective since each set is non-empty. Then it has a right inverse. Then the right inverse must send each set to an element in the set, i.e. is a choice function for A_i.)

Definition (Inverse of function)**.** An *inverse* of f is a function that is both a left inverse and a right inverse. It is written as $f^{-1} : B \to A$. It exists iff f is bijective, and is necessarily unique.

2.3 Relations

Definition (Relation)**.** A *relation* R on A specifies that some elements of A are related to some others. Formally, a relation is a subset $R \subseteq A \times A$. We write aRb iff $(a, b) \in R$.

Example. The following are examples of relations on natural numbers:

(i) aRb iff a and b have the same final digit. e.g. $(37)R(57)$.

(ii) aRb iff a divides b. e.g. $2R6$ and $2 \not R 7$.

(iii) aRb iff $a \neq b$.

(iv) aRb iff $a = b = 1$.

(v) aRb iff $|a - b| \leq 3$.

(vi) aRb iff either $a, b \geq 5$ or $a, b \leq 4$.

Again, we wish to classify different relations.

Definition (Reflexive relation)**.** A relation R is *reflexive* if

$$(\forall a)\, aRa.$$

Definition (Symmetric relation)**.** A relation R is *symmetric* iff

$$(\forall a, b)\, aRb \Leftrightarrow bRa.$$

Definition (Transitive relation)**.** A relation R is *transitive* iff

$$(\forall a, b, c)\, aRb \wedge bRc \Rightarrow aRc.$$

Example. With regards to the examples above,

Examples	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Reflexive	✓	✓	×	×	✓	✓
Symmetric	✓	×	✓	✓	✓	✓
Transitive	✓	✓	×	✓	×	✓

Definition (Equivalence relation)**.** A relation is an *equivalence relation* if it is reflexive, symmetric and transitive. e.g. (i) and (vi) in the above examples are equivalence relations.

If it is an equivalence relation, we usually write $\sim$ instead of R. As the name suggests, equivalence relations are used to describe relations that are similar to equality. For example, if we want to represent rational numbers as a pair of integers, we might have an equivalence relation defined by $(n, m) \sim (p, q)$ iff $nq = mp$, such that two pairs are equivalent if they represent the same rational number.

Example. If we consider a deck of cards, define two cards to be related if they have the same suite.

As mentioned, we like to think of things related by $\sim$ as equal. Hence we want to identify all "equal" things together and form one new object.

Definition (Equivalence class)**.** If $\sim$ is an equivalence relation, then the *equivalence class* $[x]$ is the set of all elements that are related via $\sim$ to x.

Example. In the cards example, $[8\heartsuit]$ is the set of all hearts.

Definition (Partition of set)**.** A *partition* of a set X is a collection of subsets A_α of X such that each element of X is in exactly one of A_α.

Theorem. If $\sim$ is an equivalence relation on A, then the equivalence classes of $\sim$ form a partition of A.

Proof. By reflexivity, we have $a \in [a]$. Thus the equivalence classes cover the whole set. We must now show that for all $a, b \in A$, either $[a] = [b]$ or $[a] \cap [b] = \emptyset$.

Suppose $[a] \cap [b] \neq \emptyset$. Then $\exists c \in [a] \cap [b]$. So $a \sim c, b \sim c$. By symmetry, $c \sim b$. By transitivity, we have $a \sim b$. For all $b' \in [b]$, we have $b \sim b'$. Thus by transitivity, we have $a \sim b'$. Thus $[b] \subseteq [a]$. By symmetry, $[a] \subseteq [b]$ and $[a] = [b]$. □

On the other hand, each partition defines an equivalence relation in which two elements are related iff they are in the same partition. Thus partitions and equivalence relations are "the same thing".

Definition (Quotient map)**.** The *quotient map* q maps each element in A to the equivalence class containing a, i.e. $a \mapsto [a]$. e.g. $q(8\heartsuit) = \{\heartsuit\}$.

3 Division

If you think you already know how to divide, well perhaps you are right. However, in this chapter, we will define these notions formally and properly prove things we already know (plus maybe some new things).

3.1 Euclid's Algorithm

Definition (Factor of integers)**.** Given $a, b \in \mathbb{Z}$, we say a *divides* b, a is a *factor* of b or $a \mid b$ if $(\exists c \in \mathbb{Z})\, b = ac$. For any b, ± 1 and $\pm b$ are always factors of b. The other factors are called *proper factors*.

Theorem (Division Algorithm)**.** Given $a, b \in \mathbb{Z}$, $b \neq 0$, there are unique $q, r \in \mathbb{Z}$ with $a = qb + r$ and $0 \leq r < b$.

Despite the name, the division algorithm is not an algorithm in the usual sense. Instead, it merely states that you can divide. Even the proof does not specify a (non-brute force) way of how to divide.

Proof. Choose $q = \max\{q : qb \leq a\}$. This maximum exists because the set of all q such that $qb \leq a$ is finite. Now write $r = a - qb$. We have $0 \leq r < b$ and thus q and r are found.

To show that they are unique, suppose that $a = qb + r = q'b + r'$. We have $(q - q')b = (r' - r)$. Since both r and r' are between 0 and b, we have $-b < r - r' < b$. However, $r' - r$ is a multiple of b. Thus $q - q' = r' - r = 0$. Consequently, $q = q'$ and $r = r'$. □

Definition (Common factor of integers)**.** A *common factor* of a and b is a number $c \in \mathbb{Z}$ such that $c \mid a$ and $c \mid b$.

Definition (Highest common factor/greatest common divisor)**.** The *highest common factor* or *greatest common divisor* of two numbers $a, b \in \mathbb{N}$ is a number $d \in \mathbb{N}$ such that d is a common factor of a and b, and if c is also a common factor, $c \mid d$.

Clearly if the hcf exists, it must be the largest common factor, since all other common factors divide it, and thus necessarily unique.

You might think reasonably it is more natural to define $\text{hcf}(a, b)$ to be the largest common factor. Then show that it has the property that all common factors divide it. But the above definition is superior because it does not require a prior ordering of the natural numbers (and can be extended to any ring even if they are not ordered, as we will do in IB Groups, Rings and Modules).

Notation. We write $d = \text{hcf}(a, b) = \gcd(a, b) = (a, b)$.

Here we use (a, b) to stand for a number, and has nothing to do with an ordered pair.

Proposition. If $c \mid a$ and $c \mid b$, $c \mid (ua + vb)$ for all $u, v \in \mathbb{Z}$.

Proof. By definition, we have $a = kc$ and $b = lc$. Then $ua + vb = ukc + vlc = (uk + vl)c$. So $c \mid (ua + vb)$. □

Theorem. Let $a, b \in \mathbb{N}$. Then (a, b) exists.

Proof. Let $S = \{ua + vb : u, v \in \mathbb{Z}\}$ be the set of all linear combinations of a, b. Let d be the smallest positive member of S. Say $d = xa + yb$. Hence if $c \mid a, c \mid b$, then $c \mid d$. So we need to show that $d \mid a$ and $d \mid b$, and thus $d = (a, b)$.

By the division algorithm, there exist numbers $q, r \in \mathbb{Z}$ with $a = qd + r$ with $0 \leq r < d$. Then $r = a - qd = a(1 - qx) - qyb$. Therefore r is a linear combination of a and b. Since d is the smallest positive member of S and $0 \leq r < d$, we have $r = 0$ and thus $d \mid a$. Similarly, we can show that $d \mid b$. □

Corollary. (from the proof) Let $d = (a, b)$, then d is the smallest positive linear combination of a and b.

Corollary (Bézout's identity)**.** Let $a, b \in \mathbb{N}$ and $c \in \mathbb{Z}$. Then there exists $u, v \in \mathbb{Z}$ with $c = ua + vb$ iff $(a, b) \mid c$.

Proof. ($\Rightarrow$) Let $d = (a, b)$. If c is a linear combination of a and b, then $d \mid c$ because $d \mid a$ and $d \mid b$.

($\Leftarrow$) Suppose that $d \mid c$. Let $d = xa + yb$ and $c = kd$. Then $c = (kx)a + (ky)b$. Thus c is a linear combination of a and b. □

Note that the proof that (a, b) exists is existential, not constructive. How can we actually find d, and how can we find x, y such that $d = xa + yb$?

While it might be easy to simply inspect d for small numbers, how would you find common factors of, say, 4931 and 3795? We cannot use primes because (a) prime factorization is hard; and (b) primes are not yet defined.

You might spot if $c \mid 4931$ and $c \mid 3795$, then $c \mid (4931 - 3795) = 1136$. The process is also reversible — if $c \mid 1136$ and $c \mid 3795$, then $c \mid (1136 + 3795) = 4931$. Thus the problem is equivalent to finding common factors of 3795 and 1136. The process can be repeated until we have small numbers.

Proposition (Euclid's Algorithm)**.** If we continuously break down a and b by the following procedure:

$$\begin{aligned} a &= q_1 b + r_1 \\ b &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \\ &\vdots \\ r_{n-2} &= q_n r_{n-1} \end{aligned}$$

then the highest common factor is r_{n-1}.

Proof. We have (common factors of a, b) = (common factors of b, r_1) = (common factors of r_1, r_2) = $\cdots$ = (factors of r_{n-1}). □

This gives an alternative proof that hcfs exist.

How efficient is this algorithm? For every step, we have $a \geq b + r_1 > 2r_1$. Thus every two steps, the number on the left goes down by at least half. Hence the number of digits goes down every 8 steps. Thus the time needed is $\leq 8\times$ number of digits and has time complexity $O(\log b)$.

Example. Suppose $a = 57$ and $b = 42$.

common factors of 57 and 42	$57 = 1 \times 42 + 15$
= common factors of 42 and 15	$42 = 2 \times 15 + 12$
= common factors of 15 and 12	$15 = 1 \times 12 + 3$
= common factors of 12 and 3	$12 = 4 \times 3 + 0$
= common factors of 3 and 0	
= factors of 3.	

So the hcf is 3.

By reversing Euclid's Algorithm, we can find the hcf of two numbers as a linear combination of a and b.

Example. Consider 57 and 21.

$$\begin{aligned} 57 &= 2 \times 21 + 15 \\ 21 &= 1 \times 15 + 6 \\ 15 &= 2 \times 6 + 3 \\ 6 &= 2 \times 3 \end{aligned}$$

In the opposite direction, we have

$$\begin{aligned} 3 &= 15 - 2 \times 6 \\ &= 15 - 2 \times (21 - 15) \\ &= 3 \times 15 - 2 \times 21 \\ &= 3 \times (57 - 2 \times 21) - 2 \times 21 \\ &= 3 \times 57 - 8 \times 21 \end{aligned}$$

This gives an alternative constructive proof of Bézout's identity. Moreover, it gives us a quick way of expressing $(a, b) = ax + by$. However, this algorithm requires storing the whole process of Euclid's Algorithm and is not efficient space-wise.

To achieve higher space efficiency, we attempt to find a recurrence relation for the coefficients A_j, B_j such that $a \times B_j - b \times A_j = (-1)^j r_j$. The possible factor of -1 is there just so that the recurrence relation will look nicer. Suppose that this is satisfied for all indices less than j. Then we have

$$\begin{aligned} (-1)^j r_j &= (-1)^j (r_{j-2} - q_j r_{j-1}) \\ &= (-1)^{j-2} r_{j-2} + q_j (-1)^{j-1} r_{j-1} \\ &= a(B_{j-2} + q_j B_{j-1}) - b(A_{j-2} + q_j A_{j-1}). \end{aligned}$$

Hence we can obtain the following recurrence relation:

$$\begin{aligned} A_j &= q_j A_{j-1} + A_{j-2} \\ B_j &= q_j B_{j-1} + B_{j-2} \end{aligned}$$

with

$$a \times B_j - b \times A_j = (-1)^j r_j.$$

In particular, $a \times B_{n-1} - b \times A_{n-1} = (-1)^{n-1} r_{n-1} = (a, b)$.

Also, by an easy induction, $A_j B_{j-1} - B_j A_{j-1} = (-1)^j$. So $(A_j, B_j) = 1$.

These coefficients also play another role. We can put the Euclid's Algorithm's equations in the following form:

$$\begin{aligned}\frac{57}{21} &= 2 + \frac{15}{21}\\ \frac{21}{15} &= 1 + \frac{6}{15}\\ \frac{15}{6} &= 2 + \frac{3}{6}\\ \frac{6}{3} &= 2\end{aligned}$$

Then we can write out the fraction $\frac{57}{21}$ in continued fraction form

$$\frac{57}{21} = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{2}}}$$

Expanding this continued fractions term by term, we can have the sequence $2, 2 + \frac{1}{1} = 3$, $2 + \frac{1}{1+\frac{1}{2}} = \frac{8}{3}$. These are called the "convergents". The sequence happens to be $\frac{A_i}{B_i}$.

3.2 Primes

There are a lot of ways we can define prime numbers in $\mathbb{N}$. The definition we will use is the following:

Definition (Prime number)**.** $p \in \mathbb{N}$ is a *prime* if $p > 1$ and the only factors of p (in $\mathbb{Z}$) are ± 1 and $\pm p$.

In this chapter, the objective is to prove things we already know and think are obvious.

Theorem. Every number can be written as a product of primes.

Proof. If $n \in \mathbb{N}$ is not a prime itself, then by definition $n = ab$. If either a or b is not prime, then that number can be written as a product, say $b = cd$. Then $n = acd$ and so on. Since these numbers are getting smaller, and the process will stop when they are all prime. □

In the proof, we handwaved a bit when we said "and so on". We will later come up with the principle of (strong) induction that rigorously justifies this. This is the case for many proofs we will have here.

Theorem. There are infinitely many primes.

Proof. (Euclid's proof) Suppose there are finitely many primes, say $p_1, p_2 \cdots p_n$. Then $N = p_1 p_2 \cdots p_n + 1$ is divisible by none of the primes. Otherwise, $p_j \mid (N - p_1 p_2 \cdots p_n)$, i.e. $p_j \mid 1$, which is impossible. However, N is a product of primes, so there must be primes not amongst $p_1, p_2 \cdots p_n$. □

Proof. (Erdös 1930) Suppose that there are finitely many primes, $p_1, p_2 \cdots p_k$. Consider all numbers that are the products of these primes, i.e. $p_1^{j_1} p_2^{j_2} \cdots p_k^{j_k}$, where $j_i \geq 0$. Factor out all squares to obtain the form $m^2 p_1^{i_1} p_2^{i_2} \cdots p_k^{i_k}$, where $m \in \mathbb{N}$ and $i_j = 0$ or 1.

Let $N \in \mathbb{N}$. Given any number $x \leq N$, when put in the above form, we have $m \leq \sqrt{N}$. So there are at most $\sqrt{N}$ possible values of m. For each m, there are 2^k numbers of the form $m^2 p_1^{i_1} p_2^{i_2} \cdots p_k^{i_k}$. So there are only $\sqrt{N} \times 2^k$ possible values of x of this kind.

Now pick $N \geq 4^k$. Then $N > \sqrt{N} \times 2^k$. So there must be a number $\leq N$ not of this form, i.e. it has a prime factor not in this list. □

Historically, many people have came up with "new" proofs that there are infinitely many primes. However, most of these proofs were just Euclid's proof in disguise. Erdös' proof is genuinely a new proof. For example, Euclid's proof comes up with a *particular* number N, and says *all* its factors are not in the list of primes. On the other hand, Erdös' proof says that there is *some* number, which we don't know, with *at least one* factor not in the list.

Also, the proofs give different bounds on when we should expect to see the kth prime. For example, Euclid tells us that the kth prime must be less than 2^{2^k}, while Erdös tells us it is less than 4^k.

Theorem. If $a \mid bc$ and $(a, b) = 1$, then $a \mid c$.

Proof. From Euclid's algorithm, there exist integers $u, v \in \mathbb{Z}$ such that $ua+vb = 1$. So multiplying by c, we have $uac + vbc = c$. Since $a \mid bc$, a |LHS. So $a \mid c$. □

Definition (Coprime numbers)**.** We say a, b are *coprime* if $(a, b) = 1$.

Corollary. If p is a prime and $p \mid ab$, then $p \mid a$ or $p \mid b$. (True for all p, a, b)

Proof. We know that $(p, a) = p$ or 1 because p is a prime. If $(p, a) = p$, then $p \mid a$. Otherwise, $(p, a) = 1$ and $p \mid b$ by the theorem above. □

Corollary. If p is a prime and $p \mid n_1 n_2 \cdots n_i$, then $p \mid n_i$ for some i.

Note that when we defined primes, we defined it in terms of factors of p. This corollary is the opposite — it is about how p behaves as a factor of other numbers.

Theorem (Fundamental Theorem of Arithmetic)**.** Every natural number is expressible as a product of primes in exactly one way. In particular, if $p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_l$, where p_i, q_i are primes but not necessarily distinct, then $k = l$. $q_1, \cdots q_l$ are $p_1, \cdots p_k$ in some order.

Proof. Since we already showed that there is at least one way above, we only need to show uniqueness.

Let $p_1 \cdots p_k = q_1 \cdots q_l$. We know that $p_1 \mid q_1 \cdots q_l$. Then $p_1 \mid q_1(q_2 q_3 \cdots q_l)$. Thus $p_1 \mid q_i$ for some i. wlog assume $i = 1$. Then $p_1 = q_1$ since both are primes. Thus $p_2 p_3 \cdots p_k = q_2 q_3 \cdots q_l$. Likewise, we have $p_2 = q_2, \cdots$ and so on. □

Corollary. If $a = p_1^{i_1} p_2^{i_2} \cdots p_r^{i_r}$ and $b = p_1^{j_1} p_2^{j_2} \cdots p_r^{j_r}$, where p_i are distinct primes (exponents can be zero). Then $(a, b) = \prod p_k^{\min\{i_k, j_k\}}$. Likewise, $\text{lcm}(a, b) = \prod p_k^{\max\{i_k, j_k\}}$. We have $\text{hcf}(a, b) \times \text{lcm}(a, b) = ab$.

However, this is not an efficient way to calculate (a, b), since prime factorization is very hard.

Note that this is a property peculiar to natural numbers. There are "arithmetical systems" (permitting addition, multiplication and subtraction) where factorization is not unique, e.g. even numbers.

Example. The following systems have no prime unique factorization

(i) Even numbers. "Primes" are twice of odd numbers. So 6 is a prime (NOT divisible by 2!) while 8 is not. We have $60 = 2 \times 30 = 6 \times 10$, where $2, 6, 10, 30$ are primes. However, this example is not "proper" since there is no identity element. (i.e. not a ring)

(ii) Consider $\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} : a, b \in \mathbb{Z}\}$. We have $6 = 2 \times 3 = (1-\sqrt{-5})(1+\sqrt{-5})$. It can be shown that these are primes (see IB Groups, Rings and Modules).

Exercise: Where does the proof of the Fundamental Theorem of Arithmetic fail in these examples?

4 Counting and integers

This chapter exists because experience shows that mathematicians do not know how to count.

4.1 Basic counting

A useful theorem is the pigeonhole principle.

Theorem (Pigeonhole Principle)**.** If we put $mn + 1$ pigeons into n pigeonholes, then some pigeonhole has at least $m + 1$ pigeons.

Example. In Cambridge, there are 2 people who have the same number of hairs.

Another useful tool for counting is the indicator function.

Definition (Indicator function/characteristic function)**.** Let X be a set. For each $A \subseteq X$, the *indicator function* or *characteristic function* of A is the function $i_A : X \to \{0, 1\}$ with $i_A(x) = 1$ if $x \in A$, 0 otherwise. It is sometimes written as χ_A.

Proposition.

(i) $i_A = i_B \Leftrightarrow A = B$

(ii) $i_{A\cap B} = i_A i_B$

(iii) $i_{\bar{A}} = 1 - i_A$

(iv) $i_{A\cup B} = 1 - i_{\overline{A\cup B}} = 1 - i_{\bar{A}\cap\bar{B}} = 1 - i_{\bar{A}} i_{\bar{B}} = 1 - (1 - i_A)(1 - i_B) = i_A + i_B - i_{A\cap B}$.

(v) $i_{A\setminus B} = i_{A\cap\bar{B}} = i_A i_{\bar{B}} = i_A(1 - i_B) = i_A - i_{A\cap B}$

Example. We can use the indicator function to prove certain properties about sets:

(i) Proof that $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$:

$$\begin{aligned} i_{A\cap(B\cup C)} &= i_A i_{B\cup C} \\ &= i_A(i_B + i_C - i_B i_C) \\ &= i_A i_B + i_A i_C - i_A i_B i_C \\ i_{(A\cap B)\cup(A\cap C)} &= i_{A\cap B} + i_{A\cap C} - i_{A\cap C} i_{A\cap B} \\ &= i_A i_B + i_A i_C - i_A i_C i_A i_B \\ &= i_A i_B + i_A i_C - i_A i_B i_C \end{aligned}$$

Therefore $i_{A\cap(B\cup C)} = i_{(A\cap B)\cup(A\cap C)}$ and thus $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Note that $i_A = i_A^2$ since $i_A = 0$ or 1, and $0^2 = 0$ and $1^2 = 1$.

(ii) Proof that the symmetric difference is associative: Observe that $i_{A\Delta B} \equiv i_A + i_B \pmod 2$. Thus $i_{(A\Delta B)\Delta C} = i_{A\Delta(B\Delta C)} \equiv i_A + i_B + i_C \pmod 2$.

Indicator functions are handy for computing the sizes of finite sets because if $A \subseteq X$, then $|A| = \sum_{x \in X} i_A(x)$.

Proposition. $|A \cup B| = |A| + |B| - |A \cap B|$

Proof.

$$\begin{aligned} |A \cup B| &= \sum_{x \in X} i_{A(x) \cup B(x)} \\ &= \sum (i_A(x) + i_B(x) - i_{A \cap B}(x)) \\ &= \sum i_A(x) + \sum i_B(x) - \sum i_{A \cap B}(x) \\ &= |A| + |B| - |A \cap B| \end{aligned}$$ □

More importantly, we will use indicator functions to prove a powerful result.

Theorem (Inclusion-Exclusion Principle)**.** Let A_i be subsets of a finite set X, for $1 \leq i \leq n$. Then

$$|\bar{A}_1 \cap \cdots \cap \bar{A}_n| = |X| - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| - \cdots + (-1)^n |A_1 \cap \cdots A_n|.$$

Equivalently,

$$|A_1 \cup \cdots \cup A_n| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \cdots + (-1)^{n-1} |A_1 \cap \cdots \cap A_n|.$$

The two forms are equivalent since $|A_1 \cup \cdots \cup A_n| = |X| - |\bar{A}_1 \cap \cdots \cap \bar{A}_n|$.

Proof. Using indicator functions,

$$\begin{aligned} i_{\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n} &= \prod_j i_{\bar{A}_j} \\ &= \prod_j (1 - i_{A_j}) \\ &= 1 - \sum_i i_{A_i} + \sum_{i<j} i_{A_i} i_{A_j} - \cdots + (-1)^n i_{A_1} i_{A_2} \cdots i_{A_n} \\ &= 1 - \sum_i i_{A_i} + \sum_{i<j} i_{A_i \cap A_j} - \cdots + (-1)^n i_{A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n} \end{aligned}$$

Thus

$$\begin{aligned} |\bar{A}_1 \cap \cdots \cap \bar{A}_n| &= \sum_{x \in X} i_{\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n}(x) \\ &= \sum_x 1 - \sum_i \sum_x i_{A_i}(x) + \sum_{i<j} \sum_x i_{A_i \cap A_j}(x) - \cdots \\ &+ \sum_x (-1)^n i_{A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n}(x) \\ &= |X| - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| \\ &- \sum_{i<j<k} |A_i \cap A_j \cap A_k| + \cdots + (-1)^n |A_1 \cap A_2 \cap \cdots A_n| \end{aligned}$$ □

Example. How many numbers ≤ 200 are coprime to 110?

Let $X = \{1, \cdots 200\}$, and $A_1 = \{x : 2 \mid x\}$, $A_2 = \{x : 5 \mid x\}$, $A_3 = \{x : 11 \mid x\}$. We know that

$$\begin{aligned} |A_1| &= \lfloor 200/2 \rfloor = 100 \\ |A_2| &= \lfloor 200/5 \rfloor = 40 \\ |A_3| &= \lfloor 200/11 \rfloor = 18 \\ |A_1 \cap A_2| &= \lfloor 200/10 \rfloor = 20 \\ |A_1 \cap A_3| &= \lfloor 200/22 \rfloor = 9 \\ |A_2 \cap A_3| &= \lfloor 200/55 \rfloor = 3 \\ |A_1 \cap A_2 \cap A_3| &= \lfloor 200/110 \rfloor = 1 \end{aligned}$$

Then the answer is $200 - 100 - 40 - 18 + 20 + 9 + 3 - 1 = 73$.

4.2 Combinations

"Combinations" is about counting the ways we can pick things without regards to order. We can formulate these problems in the language of sets: given a set X, how many subsets of X are there that satisfy some particular properties? For example, if we want to pick 3 people from 10, we can let X be the set of the 10 people. Then the number of ways of picking the 3 people is the number of subsets of size three.

Example. How many subsets of $\{1, 2, \cdots n\}$ are there? There are $2 \times 2 \times \cdots \times 2 = 2^n$. Since for each subset, every element is either in or out of the subset, and there are two choices for each element. Equivalently, there are 2^n possible indicator functions, i.e. functions $\{1, 2, 3, \cdots, n\} \to \{0, 1\}$.

Definition (Combination $\binom{n}{r}$)**.** The number of subsets of $\{1, 2, 3, \cdots, n\}$ of size r is denoted by $\binom{n}{r}$. The symbol is pronounced as "n choose r".

This is the *definition* of $\binom{n}{r}$. This does not in any way specify how we can actually calculate the value of $\binom{n}{r}$.

Proposition. By definition,

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n$$

Theorem (Binomial theorem)**.** For $n \in \mathbb{N}$ with $a, b \in \mathbb{R}$, we have

$$(a+b)^n = \binom{n}{0} a^n b^0 + \binom{n}{1} a^{n-1} b^1 + \cdots + \binom{n}{r} a^{n-r} b^r + \cdots + \binom{n}{n} a^0 b^n$$

Proof. We have $(a+b)^n = (a+b)(a+b)\cdots(a+b)$. When we expand the product, we get all terms attained by choosing b from some brackets, a from the rest. The term $a^{n-r}b^r$ comes from choosing b from r brackets, a from the rest, and there are $\binom{n}{r}$ ways to make such a choice. □

This theorem is not immediately useful since we do not know the value of $\binom{n}{r}$!

Because of this theorem, $\binom{n}{r}$ is sometimes called a "binomial coefficient".

Proposition.

(i) $\binom{n}{r} = \binom{n}{n-r}$. This is because choosing r things to keep is the same as choosing $n-r$ things to throw away.

(ii) $\binom{n}{r-1} + \binom{n}{r} = \binom{n+1}{r}$ (Pascal's identity) The RHS counts the number of ways to choose a team of r players from $n+1$ available players, one of whom is Pietersen. If Pietersen is chosen, there are $\binom{n}{r-1}$ ways to choose the remaining players. Otherwise, there are $\binom{n}{r}$ ways. The total number of ways is thus $\binom{n}{r-1} + \binom{n}{r}$.

Now given that $\binom{n}{0} = \binom{n}{n} = 1$, since there is only one way to choose nothing or everything, we can construct *Pascal's triangle*:

$$\begin{array}{ccccccccccc} & & & & & 1 & & & & & \\ & & & & 1 & & 1 & & & & \\ & & & 1 & & 2 & & 1 & & & \\ & & 1 & & 3 & & 3 & & 1 & & \\ & 1 & & 4 & & 6 & & 4 & & 1 & \end{array}$$

where each number is the sum of the two numbers above it, and the rth item of the nth row is $\binom{n}{r}$ (first row is row 0).

(iii) $\binom{n}{k}\binom{k}{r} = \binom{n}{r}\binom{n-r}{k-r}$. We are counting the number of pairs of sets (Y, Z) with $|Y| = k$ and $|Z| = r$ with $Z \subseteq Y$. In the LHS, we first choose Y then choose $Z \subseteq Y$. The RHS chooses Z first and then choose the remaining $Y \setminus Z$ from $\{1, 2, \cdots n\} \setminus Z$.

(iv) $\binom{a}{r}\binom{b}{0} + \binom{a}{r-1}\binom{b}{1} + \cdots + \binom{a}{r-k}\binom{b}{k} + \cdots \binom{a}{0}\binom{b}{r} = \binom{a+b}{r}$ (Vandermonde's convolution) Suppose we have a men and b women, and we need to choose a committee of r people. The right hand side is the total number of choices. The left hand side breaks the choices up according to the number of men vs women.

Example. A greengrocer stocks n kinds of fruit. In how many ways can we choose a bag of r fruits? If we are only allowed to choose one of each kind, then the answer is $\binom{n}{r}$. But we might have $r = 4$, and we want to allow picking 2 apples, 1 plum and 1 quince. The total number of ways to choose is $\binom{n+r-1}{r}$. Why?

Each choice can be represented by a binary string of length $n + r - 1$, with r 0's and $n - 1$ 1's. The string can be constructed as follows (by example): when $n = 5$ and $r = 8$, a possible binary string 000100110010. The block of zeros corresponds to the number of each fruit chosen, and the 1s separate the choices. In the string above, we have 3 of type 1, 2 of type 2, 0 of type 3, 2 of type 4 and 1 of type 5.

Then clearly the number of possible strings is $\binom{n+r-1}{r}$.

Proposition. $\binom{n}{r} = \dfrac{n!}{(n-r)!r!}$.

Proof. There are $n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}$ ways to choose r elements in order. Each choice of subsets is chosen this way in $r!$ orders, so the number of subsets is $\frac{n!}{(n-r)!r!}$. □

We might write $x^{\underline{r}}$ for the polynomial $x(x-1)\cdots(x-r+1)$. We call this "x to the r falling". We can write $\binom{n}{r} = \frac{n^{\underline{r}}}{r!}$. Multiplying Vandermonde by $r!$, we obtain the "falling binomial theorem"

$$\binom{r}{0}a^{\underline{r}}b^{\underline{0}} + \binom{r}{1}a^{\underline{r-1}}b^{\underline{1}} + \cdots + \binom{r}{r}a^{\underline{0}}b^{\underline{r}} = (a+b)^{\underline{r}}.$$

Example. A bank prepares a letter for each of its n customers, saying how much it cares. (Each of these letters costs the customer £40) There are $n!$ ways to put the letters in the envelopes. In how many ways can this be done so that no one gets the right letter (i.e. how many *derangements* are there of n elements)?

We let X be the set of all envelopings (permutation of n). $|X| = n!$. For each i, let $A_i = \{x \in X : x \text{ assigns the correct letter to customer } i\}$. We want to know $|\bigcap_i \bar{A}_i|$. We know that $|A_i| = (n-1)!$ since i's letter gets in i's envelopes and all others can be placed randomly. We have $|A_i \cap A_j| = (n-2)!$ as well. Similarly, $|A_i \cap A_j \cap A_k| = (n-3)!$.

By the inclusion-exclusion formula, we have

$$\begin{aligned}\left|\bigcap_i \bar{A}_i\right| &= |X| - \sum |A_i| + \sum |A_i \cap A_j| + \cdots \\ &= n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \cdots \\ &= n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - \cdots + \frac{(-1)^n}{n!}\right) \\ &\approx n!e^{-1}\end{aligned}$$

4.3 Well-ordering and induction

Several proofs so far involved "take the least integer such that", e.g. division algorithm; or involved a sequence of moves "and so on..." e.g. Euclid's algorithm, every number is a product of primes. We rely on the following:

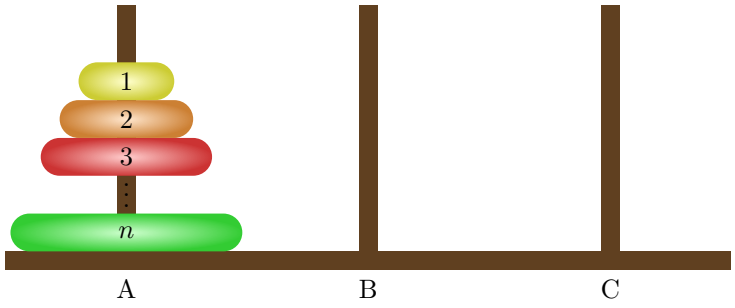
Theorem (Weak Principle of Induction)**.** Let $P(n)$ be a statement about the natural number n. Suppose that

(i) $P(1)$ is true

(ii) $(\forall n)\, P(n) \Rightarrow P(n+1)$

Then $P(n)$ is true for all $n \geq 1$.

Example (Tower of Hanoi)**.** Referring to the image below,



The objective is to move the n rings on peg A to peg B, with the constraints that you can only move one ring at a time, and you can never place a larger ring onto a smaller ring.

Now claim that this needs exactly $2^n - 1$ moves.

Let $P(n)$ be "n rings needs $2^n - 1$ moves". Note that this statement contains two assertions — (1) we can do it in $2^n - 1$ moves; (2) We can't do it in fewer.

First consider $P(1)$. We simply have to move the ring from A to B.

Suppose we have $n + 1$ rings. We can move the top n rings to peg C, then move the bottom ring to peg B, then move the n rings from C back to B. Assuming $P(n)$ is true, this needs at most $2 \times (2^n - 1) + 1 = 2^{n+1} - 1$ moves.

Can we do it in fewer moves? To succeed, we must free the bottom ring, so we must shift the top n rings to another pe.g. This needs $\geq 2^n - 1$ moves by $P(n)$. Then we need to shift the bottom ring. Then we need to shift the n smaller rings to the big one. This needs $\geq 2^n - 1$ moves by $P(n)$. So this needs $\geq 2^{n+1} - 1$ moves altogether.

So we showed that $P(n) \Rightarrow P(n+1)$ (we used $P(n)$ four times). By the WPI, $P(n)$ is true for all n.

Example. All numbers are equal. Let $P(n)$ be "if $\{a_1, \cdots a_n\}$ is a set of n numbers, then $a_1 = a_2 = \cdots a_n$". $P(1)$ is trivially true. Suppose we have $\{a_1, a_2 \cdots a_{n+1}\}$. Assuming $P(n)$, apply it to $\{a_1, a_2 \cdots a_n\}$ and $\{a_2, \cdots, a_{n+1}\}$, then $a_1 = \cdots = a_n$ and $a_2 = a_3 = \cdots = a_{n+1}$. So $a_1 = a_2 = \cdots = a_{n+1}$. Hence $P(n) \Rightarrow P(n+1)$. So $P(n)$ is true for all $n \in \mathbb{N}$.

Theorem. Inclusion-exclusion principle.

Proof. Let $P(n)$ be the statement "for any sets $A_1 \cdots A_n$", we have $|A_1 \cup \cdots \cup A_n| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \cdots \pm |A_i \cap A_2 \cap \cdots \cap A_n|$".

$P(1)$ is trivially true. $P(2)$ is also true (see above). Now given $A_1 \cdots A_{n+1}$, Let $B_i = A_i \cap A_{n+1}$ for $1 \leq i \leq n$. We apply $P(n)$ both to the A_i and B_i.

Now observe that $B_i \cap B_j = A_i \cap A_j \cap A_{n+1}$. Likewise, $B_i \cap B_j \cap B_k = A_i \cap A_j \cap A_k \cap A_{n+1}$. Now

$$\begin{aligned}
|A_1 \cup A_2 \cup \cdots \cup A_{n+1}| &= |A_1 \cup \cdots \cup A_n| + |A_{n+1}| - |(A_1 \cup \cdots \cup A_n) \cap A_{n+1}| \\
&= |A_1 \cup \cdots \cup A_n| + |A_{n+1}| - |B_1 \cup \cdots \cup B_n| \\
&= \sum_{i \leq n} |A_i| - \sum_{i<j \leq n} |A_i \cap A_j| + \cdots + |A_{n+1}| \\
&\quad - \sum_{i \leq n} |B_i| + \sum_{i<j \leq n} |B_i \cap B_j| - \cdots
\end{aligned}$$

Note $\sum_{i\le n} |B_i| = \sum_{i\le n} |A_i \cap A_{n+1}|$. So $\sum_{i<j\le n} |A_i \cap A_j| + \sum_{i\le n} |B_i| = \sum_{i<j\le n+1} |A_i \cap A_j|$, and similarly for the other terms. So

$$= \sum_{i\le n+1} |A_i| - \sum_{i<j\le n+1} |A_i \cap A_j| + \cdots$$

So $P(n) \Rightarrow P(n+1)$ for $n \ge 2$. By WPI, $P(n)$ is true for all n. □

However, WPI is not quite what we want for "every number is a product of primes". We need a different form of induction.

Theorem (Strong principle of induction)**.** Let $P(n)$ be a statement about $n \in \mathbb{N}$. Suppose that

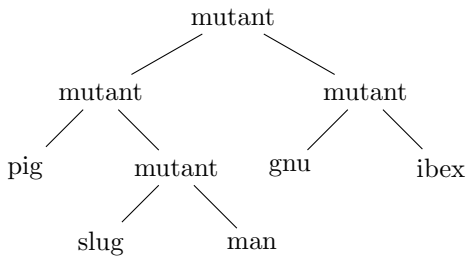
(i) $P(1)$ is true

(ii) $\forall n \in N$, if $P(k)$ is true $\forall k < n$ then $P(n)$ is true.

Then $P(n)$ is true for all $n \in N$.

Note that (i) is redundant as it follows from (ii), but we state it for clarity.

Example. "Evolutionary trees" Imagine that we have a mutant that can produce two offsprings. Each offspring is either an animal or another mutant. A possible evolutionary tree is as follows:



Let $P(n)$ be the statement $n-1$ mutants produces n animals. Given some tree with n animals, remove the top mutant to get two sub-trees, with n_1 and n_2 animals, where $n_1 + n_2 = n$. If $P(k)$ is true $\forall k < n$, then $P(n_1)$ and $P(n_2)$ are true. So the total number of mutants is $1 + (n_1 - 1) + (n_2 - 1) = n - 1$. So $P(n)$ is true. Hence by strong principle of induction, $P(n)$ is true for all n.

Theorem. The strong principle of induction is equivalent to the weak principle of induction.

Proof. Clearly the strong principle implies the weak principle since if $P(n) \Rightarrow P(n+1)$, then $(P(1) \wedge P(2) \wedge \cdots \wedge P(n)) \Rightarrow P(n+1)$.

Now show that the weak principle implies the strong principle. Suppose that $P(1)$ is true and $(\forall n)\, P(1) \wedge P(2) \wedge \cdots \wedge P(n-1) \Rightarrow P(n)$. We want to show that $P(n)$ is true for all n using the weak principle.

Let $Q(n) =$ "$P(k)$ is true $\forall k \le n$". Then $Q(1)$ is true. Suppose that $Q(n)$ is true. Then $P(1) \wedge P(2) \wedge \cdots \wedge P(n)$ is true. So $P(n+1)$ is true. Hence $Q(n+1)$ is true. By the weak principle, $Q(n)$ is true for all n. So $P(n)$ is true for all n. □

While strong and weak induction are practically equivalent, they are rather distinct conceptually. Weak induction is expressed in terms of "adding 1", while strong induction is based on the ordering of natural numbers.

It turns out that there is another statement about the natural numbers that can be stated in terms of orders. We first formally define what it means to be an order.

Definition (Partial order)**.** A *partial order* on a set is a reflexive, antisymmetric $((aRb) \wedge (bRa) \Leftrightarrow a = b)$ and transitive relation.

Example. The ordinary ordering of $\mathbb{N}$ $a \leq b$ is a partial order of $\mathbb{N}$. Also, $a \mid b$ on $\mathbb{N}$ is also a partial order.

Definition (Total order)**.** A *total order* is a partial order where $\forall a \neq b$, exactly one of aRb or bRa holds. This means that every two things must be related.

Definition (Well-ordered total order)**.** A total order is *well-ordered* if every non-empty subset has a minimal element, i.e. if $S \neq \emptyset$, then $\exists m \in S$ such that $x < m \Rightarrow x \notin S$.

Example. $\mathbb{Z}$ with the usual order is not well-ordered since the set of even integers has no minimum. The positive rationals are also not well-ordered under the usual order.

Theorem (Well-ordering principle)**.** $\mathbb{N}$ is well-ordered under the usual order, i.e. every non-empty subset of $\mathbb{N}$ has a minimal element.

Theorem. The well-ordering principle is equivalent to the strong principle of induction.

Proof. First prove that well-ordering implies strong induction. Consider a proposition $P(n)$. Suppose $P(k)$ is true $\forall k < n$ implies $P(n)$.

Assume the contrary. Consider the set $S = \{n \in \mathbb{N} : \neg P(n)\}$. Then S has a minimal element m. Since m is the minimal counterexample to P, $P(k)$ is true for all $k < m$. However, this implies that $P(m)$ is true, which is a contradiction. Therefore $P(n)$ must be true for all n.

To show that strong induction implies well-ordering, let $S \subseteq \mathbb{N}$. Suppose that S has no minimal element. We need to show that S is empty. Let $P(n)$ be the statement $n \notin S$.

Certainly $1 \notin S$, or else it will be the minimal element. So $P(1)$ is true. Suppose we know that $P(k)$ is true for all $k < n$, i.e. $k \notin S$ for all $k < n$. Now $n \notin S$, or else n will be the minimal element. So $P(n)$ is true. By strong induction, $P(n)$ is true for all n, i.e. S is empty. □

The well-ordering principle enables us to show that $P(n)$ is true as follows: if $P(n)$ fails for some n, then there is a minimal counterexample m. Then we try to show that this leads to a contradiction.

Example. Proof that every number is a product of primes by strong induction: Assume the contrary. Then there exists a minimal n that cannot be written as a product of prime (by the well-ordering principle). If n is a prime, then n is a product of primes. Otherwise, write $n = ab$, where $1 < a, b < n$. By minimality of n, both a and b are products of primes. Hence so is n. Contradiction.

Example. All numbers are interesting. Suppose that there are uninteresting numbers. Then there exists a smallest uninteresting number. Then the property of being the smallest uninteresting number is itself interesting. Contradiction.

Example. Consider a total order on $\mathbb{N} \times \mathbb{N}$ by "lexicographic" or "dictionary" order, i.e. $(a, b) \leq (c, d)$ if $a < c$ or $(a = c \wedge b \leq d)$.

The Ackermann function is a function $a : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}$ is defined by

$$a(m, n) = \begin{cases} n + 1 & \text{if } m = 0 \\ a(m - 1, 1) & \text{if } m > 0 \text{ and } n = 0 \\ a(m - 1, a(m, n - 1)) & \text{if } m > 0 \text{ and } n > 0. \end{cases}$$

We want to show that this is well-defined.

Note that $a(m, n)$ is expressed in terms of a at points $(x, y) < (m, n)$. So a is well-defined if lexicographic order is well-ordered, i.e. every non-empty subset has a minimal element (if a were not well-defined, then would be a smallest place where the definition is bad. But definition of that point is defined in terms of smaller points which are well defined).

We can see that $\mathbb{N}_0 \times \mathbb{N}_0$ is well-ordered: if $S \subseteq \mathbb{N}_0 \times \mathbb{N}_0$ is non-empty, let S_x be the set of $\{x \in \mathbb{N} : (\exists y)\, (x, y) \in S\}$, i.e. the set of all x-coordinates of S. By the well-ordering principle, S_x has a minimal element m. Then let $S_y = \{y \in \mathbb{N}_0 : (m, y) \in S\}$. Then S_y has a minimal element n. Then (m, n) is the minimal element of S.

5 Modular arithmetic

We are going to study *modular arithmetic*. In modular arithmetic, we first pick a particular natural number to be the *modulus*, say 7. Then we consider two numbers to be "equal" if their difference is a multiple of the modulus. For example, modulo 7, we will think that 3 and 10 are "the same", while 2 and 4 are different.

We will study arithmetic under this number system. Like the integers, we are allowed to add and multiply numbers. However, while in $\mathbb{Z}$, we can only divide by 1 and -1, in modular arithmetic, more numbers can be divided. For example, modulo 10, we are allowed to divide by 3.

An important application of modular arithmetic is RSA encryption. This is a widely deployed asymmetric encryption algorithm. By asymmetric, we mean that the key for encryption is different from the key for decryption. This is very useful in real life, since we can broadcast the encryption key to the world, and keep the decryption key to ourselves. This way anyone can send you encrypted messages that only you can decrypt.

5.1 Modular arithmetic

Definition (Modulo)**.** If $a, b \in \mathbb{Z}$ have the same remainder after division by m, i.e. $m \mid (a - b)$, we say a and b are *congruent modulo* m, and write

$$a \equiv b \pmod{m}$$

Example. The check digits of the ISBN (or Hong Kong ID Card Number) are calculated modulo 11.

Example. $9 \equiv 0 \pmod 3$, $11 \equiv 6 \pmod 5$.

Proposition. If $a \equiv b \pmod m$, and $d \mid m$, then $a \equiv b \pmod d$.

Proof. $a \equiv b \pmod m$ if and only if $m \mid (a - b)$, hence $d \mid (a - b)$, i.e. $a \equiv b \pmod d$. □

Observe that with m fixed, $a \equiv b \pmod m$ is an equivalence relation. The set of equivalence classes is written as $\mathbb{Z}_m$ or $\mathbb{Z}/(m\mathbb{Z})$.

Example. $\mathbb{Z}_3 = \{[0], [1], [2]\}$

Proposition. If $a \equiv b \pmod m$ and $u \equiv v \pmod m$, then $a + u \equiv b + v \pmod m$ and $au \equiv bv \pmod m$.

Proof. Since $a \equiv b \pmod m$ and $u \equiv v \pmod m$, we have $m \mid (a-b)+(u-v) = (a+u)-(b+v)$. So $a + u \equiv b + v \pmod m$

Since $a \equiv b \pmod m$ and $u \equiv v \pmod m$, we have $m \mid (a-b)u+b(u-v) = au - bv$. So $au \equiv bv \pmod m$. □

This means that we can do arithmetic modulo n. Formally, we are doing arithmetic with the congruence classes, i.e $\mathbb{Z}_m$. For example, in $\mathbb{Z}_7$, $[4] + [5] = [9] = [2]$.

Modular arithmetic can sometimes be used to show that equations have no solutions.

Example. $2a^2 + 3b^3 = 1$ has no solutions in $\mathbb{Z}$. If there were a solution, then $2a^2 \equiv 1 \pmod 3$. But $2 \cdot 0^2 \equiv 0$, $2 \cdot 1^2 \equiv 2$ and $2 \cdot 2^2 \equiv 2$. So there is no solution to the congruence, and hence none to the original equation.

Observe that all odd numbers are either $\equiv 1 \pmod 4$ or $\equiv 3 \equiv -1 \pmod 4$. So we can classify primes depending on their value modulo 4.

Theorem. There are infinitely many primes that are $\equiv -1 \pmod 4$.

Proof. Suppose not. So let $p_1, \cdots p_k$ be all primes $\equiv -1 \pmod 4$. Let $N = 4p_1p_2 \cdots p_k - 1$. Then $N \equiv -1 \pmod 4$. Now N is a product of primes, say $N = q_1q_2 \cdots q_\ell$. But $2 \nmid N$ and $p_i \nmid N$ for all i. So $q_i \equiv 1 \pmod 4$ for all i. But then that implies $N = q_1q_2 \cdots q_\ell \equiv 1 \pmod 4$, which is a contradiction. □

Example. Solve $7x \equiv 2 \pmod{10}$. Note that $3 \cdot 7 \equiv 1 \pmod{10}$. If we multiply the equation by 3, then we get $3 \cdot 7 \cdot x \equiv 3 \cdot 2 \pmod{10}$. So $x \equiv 6 \pmod{10}$. Effectively, we divided by 7.

"Division" doesn't always work for all numbers, e.g. you cannot divide by 2 mod 10. We give a name to numbers we can divide.

Definition (Unit (modular arithmetic))**.** u is a *unit* if $\exists v$ such that $uv \equiv 1 \pmod m$.

Theorem. u is a unit modulo m if and only if $(u, m) = 1$.

Proof. ($\Rightarrow$) Suppose u is a unit. Then $\exists v$ such that $uv \equiv 1 \pmod m$. Then $uv = 1 + mn$ for some n, or $uv - mn = 1$. So 1 can be written as a linear combination of u and m. So $(u, m) = 1$.

($\Leftarrow$) Suppose that $(u, m) = 1$. Then there exists a, b with $ua + mb = 1$. Thus $ua \equiv 1 \pmod m$. □

Using the above proof, we can find the inverse of a unit efficiently by Euclid's algorithm.

Corollary. If $(a, m) = 1$, then the congruence $ax \equiv b \pmod m$ has a unique solution $\pmod m$.

Proof. If $ax \equiv b \pmod m$, and $(a, m) = 1$, then $\exists a^{-1}$ such that $a^{-1}a \equiv 1 \pmod m$. So $a^{-1}ax \equiv a^{-1}b \pmod m$ and thus $x \equiv a^{-1}b \pmod m$. Finally we check that $x \equiv a^{-1}b \pmod m$ is indeed a solution: $ax \equiv aa^{-1}b \equiv b \pmod m$. □

Proposition. There is a solution to $ax \equiv b \pmod m$ if and only if $(a, m) \mid b$.

If $d = (a, m) \mid b$, then the solution is the unique solution to $\frac{a}{d}x \equiv \frac{b}{d}x \pmod{\frac{m}{d}}$

Proof. Let $d = (a, m)$. If there is a solution to $ax \equiv b \pmod m$, then $m \mid ax - b$. So $d \mid ax - b$ and $d \mid b$.

On the contrary, if $d \mid b$, we have $ax \equiv b \pmod m \Leftrightarrow ax - b = km$ for some $k \in Z$. Write $a = da'$, $b = db'$ and $m = dm'$. So $ax \equiv b \pmod m \Leftrightarrow da'x - db' = dkm' \Leftrightarrow a'x - b' = km' \Leftrightarrow a'x \equiv b' \pmod{m'}$. Note that $(a', m') = 1$ since we divided by their greatest common factor. Then this has a unique solution modulo m'. □

Example. $2x \equiv 3 \pmod 4$ has no solution since $(2, 4) = 2$ which does not divide 3.

5.2 Multiple moduli

In this section, we are concerned with multiple equations involving different moduli.

Suppose we are given $x \equiv 2 \pmod 3$ and $x \equiv 1 \pmod 4$. What is the general solution to x? We work in mod 12. Since we are given that $x \equiv 2 \pmod 3$, we know that $x \equiv 2, 5, 8$ or $11 \pmod{12}$. Similarly, since $x \equiv 1 \pmod 4$, we must have $x \equiv 1, 5$ or $9 \pmod{12}$. Combining these results, we must have $x \equiv 5 \pmod{12}$.

On the other hand, if $x \equiv 5 \pmod{12}$, then $x \equiv 5 \equiv 2 \pmod 3$ and $x \equiv 5 \equiv 1 \pmod 4$. So $x \equiv 5 \pmod{12}$ is indeed the most general solution.

In general, we have the *Chinese remainder theorem*.

Theorem (Chinese remainder theorem)**.** Let $(m, n) = 1$ and $a, b \in \mathbb{Z}$. Then there is a unique solution (modulo mn) to the simultaneous congruences

$$\begin{cases} x \equiv a \pmod m \\ x \equiv b \pmod n \end{cases},$$

i.e. $\exists x$ satisfying both and every other solution is $\equiv x \pmod{mn}$.

Proof. Since $(m, n) = 1$, $\exists u, v \in \mathbb{Z}$ with $um + vn = 1$. Then $vn \equiv 1 \pmod m$ and $um \equiv 1 \pmod n$. Put $x = umb + vna$. So $x \equiv a \pmod m$ and $x \equiv b \pmod n$.

To show it is unique, suppose both y and x are solutions to the equation. Then

$$\begin{aligned} & y \equiv a \pmod m \text{ and } y \equiv b \pmod n \\ \Leftrightarrow\ & y \equiv x \pmod m \text{ and } y \equiv x \pmod n \\ \Leftrightarrow\ & m \mid y - x \text{ and } n \mid y - x \\ \Leftrightarrow\ & mn \mid y - x \\ \Leftrightarrow\ & y \equiv x \pmod{mn} \end{aligned}$$

□

This shows a congruence (mod mn) is equivalent to one (mod n) and another (mod m).

We can easily extend this to more than two moduli by repeatedly applying this theorem.

Proposition. Given any $(m, n) = 1$, c is a unit mod mn iff c is a unit both mod m and mod n.

Proof. ($\Rightarrow$) If $\exists u$ such that $cu \equiv 1 \pmod{mn}$, then $cu \equiv 1 \pmod m$ and $cu \equiv 1 \pmod n$. So c is a unit mod m and n.

($\Leftarrow$) Suppose there exists u, v such that $cu \equiv 1 \pmod m$ and $cv \equiv 1 \pmod n$. Then by CRT, $\exists w$ with $w \equiv u \pmod m$ and $w \equiv v \pmod n$. Then $cw \equiv cu \equiv 1 \pmod m$ and $cw \equiv cv \equiv 1 \pmod n$.

But we know that $1 \equiv 1 \pmod m$ and $1 \equiv 1 \pmod n$. So 1 is a solution to $cw \equiv 1 \pmod m$, $cw \equiv 1 \pmod n$. By the "uniqueness" part of the Chinese remainder theorem, we must have $cw \equiv 1 \pmod{mn}$. □

Definition (Euler's totient function)**.** We denote by $\phi(m)$ the number of integers a, $0 \leq a \leq m$, such that $(a, m) = 1$, i.e. a is a unit mod m. Note $\phi(1) = 1$.

Proposition.

(i) $\phi(mn) = \phi(m)\phi(n)$ if $(m, n) = 1$, i.e. ϕ is multiplicative.

(ii) If p is a prime, $\phi(p) = p - 1$

(iii) If p is a prime, $\phi(p^k) = p^k - p^{k-1} = p^k(1 - 1/p)$

(iv) $\phi(m) = m\prod_{p|m}(1 - 1/p)$.

Proof. We will only prove (iv). In fact, we will prove it twice.

(i) Suppose $m = p_1^{k_1}p_2^{k_2}\cdots p_\ell^{k_\ell}$. Then

$$\begin{aligned}\phi(m) &= \phi(p_1^{k_1})\phi(p_2^{k_2})\cdots\phi(p_\ell^{k_\ell})\\ &= p_1^{k_1}(1 - 1/p_1)p_2^{k_2}(1 - 1/p_2)\cdots p_\ell^{k_\ell}(1 - 1/p_\ell)\\ &= m\prod_{p|m}(1 - 1/p)\end{aligned}$$

(ii) Let $m = p_1^{k_1}p_2^{k_2}\cdots p_\ell^{k_\ell}$. Let $X = \{0, \cdots m - 1\}$. Let $A_j = \{x \in X : p_j \mid x\}$. Then $|X| = m$, $|A_j| = m/p_j$, $|A_i \cap A_j| = m/(p_ip_j)$ etc. So $\phi(m) = |\bar{A}_1 \cap \bar{A}_2 \cap \cdots \bar{A}_\ell| = m\prod_{p|m}(1 - 1/p)$. □

Example. $\phi(60) = 60(1 - 1/2)(1 - 1/3)(1 - 1/5) = 16$.

If a, b are both units (mod m), then so is ab, for if $au \equiv 1$ and $bv \equiv 1$, then $(ab)(uv) \equiv 1$. So the units form a multiplicative group of size $\phi(m)$.

5.3 Prime moduli

Modular arithmetic has some nice properties when the modulus is a prime number.

Theorem (Wilson's theorem)**.** $(p - 1)! \equiv -1 \pmod p$ if p is a prime.

Proof. If p is a prime, then $1, 2, \cdots, p - 1$ are units. Among these, we can pair each number up with its inverse (e.g. 3 with 4 in modulo 11). The only elements that cannot be paired with a different number are 1 and -1, who are self-inverses, as show below:

$$\begin{aligned}&x^2 \equiv 1 \pmod p\\ \Leftrightarrow\ &p \mid (x^2 - 1)\\ \Leftrightarrow\ &p \mid (x - 1)(x + 1)\\ \Leftrightarrow\ &p \mid x - 1 \text{ or } p \mid x + 1\\ \Leftrightarrow\ &x \equiv \pm 1 \pmod p\end{aligned}$$

Now $(p - 1)!$ is a product of $(p - 3)/2$ inverse pairs together with 1 and -1. So the product is -1. □

Theorem (Fermat's little theorem)**.** Let p be a prime. Then $a^p \equiv a \pmod p$ for all $a \in \mathbb{Z}$. Equivalently, $a^{p-1} \equiv 1 \pmod p$ if $a \not\equiv 0 \pmod p$.

Proof. Two proofs are offered:

(i) The numbers $\{1, 2, \cdots p-1\}$ are units modulo p and form a group of order $p-1$. So $a^{p-1} \equiv 1$ by Lagrange's theorem.

(ii) If $a \not\equiv 0$, then a is a unit. So $ax \equiv ay$ iff $x \equiv y$. Then $a, 2a, 3a, \cdots (p-1)a$ are distinct mod p. So they are congruent to $1, 2, \cdots p-1$ in some order. Hence $a \cdot 2a \cdots 3a \cdots (p-1)a \equiv 1 \cdot 2 \cdot 3 \cdots (p-1)$. So $a^{p-1}(p-1)! \equiv (p-1)!$. So $a^{p-1} \equiv 1 \pmod{p}$. □

Neither Wilson nor Fermat's theorem hold if the modulus is non-prime. However, Fermat's theorem can be generalized:

Theorem (Fermat-Euler Theorem)**.** Let a, m be coprime. Then

$$a^{\phi(m)} \equiv 1 \pmod{m}.$$

Proof. Lagrange's theorem: The units mod m form a group of size $\phi(m)$.

Alternatively, let $U = \{x \in \mathbb{N} : 0 < x < m, (x, m) = 1\}$. These are the $\phi(m)$ units. Since a is a unit, $ax \equiv ay \pmod{m}$ only if $x \equiv y \pmod{m}$. So if $U = \{u_1, u_2, \cdots, u_{\phi(m)}\}$, then $\{au_1, au_2, \cdots au_{\phi(m)}\}$ are distinct and are units. So they must be $u_1, \cdots u_{\phi(m)}$ in some order. Then $au_1 au_2 \cdots au_{\phi(m)} \equiv u_1 u_2 \cdots u_{\phi(m)}$. So $a^{\phi(m)} z \equiv z$, where $z = u_1 u_2 \cdots u_{\phi(m)}$. Since z is a unit, we can multiply by its inverse and obtain $a^{\phi(m)} \equiv 1$. □

Definition (Quadratic residues)**.** The *quadratic residues* are the "squares" mod p, i.e. $1^2, 2^2, \cdots, (p-1)^2$.

Note that if $a^2 \equiv b^2 \pmod{p}$, then $p \mid a^2 - b^2 = (a-b)(a+b)$. Then $p \mid a-b$ or $p \mid a+b$. So $a \equiv \pm b \pmod{p}$. Thus every square is a square of exactly two numbers.

Example. If $p = 7$, then $1^2 \equiv 6^2 \equiv 1$, $2^2 \equiv 5^2 \equiv 4$, $3^2 \equiv 4^2 \equiv 2$. So $1, 2, 4$ are quadratic residues. $3, 5, 6$ are not.

Proposition. If p is an odd prime, then -1 is a quadratic residue if and only if $p \equiv 1 \pmod{4}$.

Proof. If $p \equiv 1 \pmod{4}$, say $p = 4k+1$, then by Wilson's theorem, $-1 \equiv (p-1)! \equiv 1 \cdot 2 \cdots \left(\frac{p-1}{2}\right)\left(-\frac{p-1}{2}\right) \cdots (-2)(-1) \equiv (-1)^{\frac{p-1}{2}} \left(\frac{p-1}{2}\right)!^2 = (-1)^{2k}(2k!)^2 \equiv (2k!)^2$. So -1 is a quadratic residue.

When $p \equiv -1 \pmod{4}$, i.e. $p = 4k+3$, suppose -1 is a square, i.e. $-1 \equiv z^2$. Then by Fermat's little theorem, $1 \equiv z^{p-1} \equiv z^{4k+2} \equiv (z^2)^{2k+1} \equiv (-1)^{2k+1} \equiv -1$. Contradiction. □

Proposition. (Unproven) A prime p is the sum of two squares if and only if $p \equiv 1 \pmod{4}$.

Proposition. There are infinitely many primes $\equiv 1 \pmod{4}$.

Proof. Suppose not, and $p_1, \cdots p_k$ are all the primes $\equiv 1 \pmod{4}$. Let $N = (2p_1 \cdots p_k)^2 + 1$. Then N is not divisible by 2 or $p_1, \cdots, p_k$. Let q be a prime $q \mid N$. Then $q \equiv -1 \pmod{4}$. Then $N \equiv 0 \pmod{q}$ and hence $(2p_1 \cdots p_k)^2 + 1 \equiv 0 \pmod{q}$, i.e. $(2p_1 \cdots p_k)^2 \equiv -1 \pmod{q}$. So -1 is a quadratic residue mod q, which is a contradiction since $q \equiv -1 \pmod{4}$. □

Proposition. Let $p = 4k + 3$ be a prime. Then if a is a quadratic residue, i.e. $a \equiv z^2 \pmod{p}$ for some z, then $z = \pm a^{k+1}$.

Proof. By Fermat's little theorem, $a^{2k+1} \equiv z^{4k+2} \equiv z^{p-1} \equiv 1$. If we multiply by a, then $a^{2k+2} \equiv a \pmod{p}$. So $(\pm a^{k+1})^2 \equiv a \pmod{p}$. □

This allows us to take square roots efficiently. This efficiency requires an effective way of computing powers of a efficiently. This can be done by repeated squaring. For example, to find a^{37}, we can calculate this by $a^{37} = a^{32}a^4a^1 = ((((a^2)^2)^2)^2)^2 \cdot (a^2)^2 \cdot a$. Thus calculation of a^n has time complexity $O(\log n)$, as opposed to $O(n)$ if you take powers manually.

Suppose a is a square mod n, where $n = pq$ and p, q are distinct primes. Then a is a square mod p and a square mod q. So there exists some s with $(\pm s)^2 \equiv a \pmod{p}$ and some t with $(\pm t)^2 \equiv a \pmod{q}$. By the Chinese remainder theorem, we can find a unique solution of each case, so we get 4 square roots of a modulo n.

5.4 Public-key (asymmetric) cryptography

Tossing a coin over a phone

Suppose we have Alice and Bob who wish to toss a coin fairly over the phone. Alice chooses two 100 digit primes with $p, q \equiv 3 \pmod{4}$. Then she tells Bob the product $n = pq$. Bob picks a number u coprime to n, computes $a \equiv u^2 \pmod{n}$ and tells Alice the value of a.

Alice can compute the square roots of a by the above algorithm ($O(\log n)$), obtain $\pm u, \pm v$ and tells Bob one of these pairs.

Now if Alice picks $\pm u$, Bob says "you win". Otherwise, Bob says "you lose".

Can Bob cheat? If he says "you lose" when Alice says $\pm u$, Bob must produce the other pair $\pm v$, but he can't know $\pm v$ without factorizing n. (If he knows $\pm u$ and $\pm v$, then $u^2 \equiv v^2 \pmod{n}$, then $n \mid (u - v)(u + v)$. But $n \nmid (u - v)$ and $n \nmid (u + v)$. So $p \mid (u - v)$ and $q \mid (u + v)$. Then $p = (n, u - v)$ and $q = (n, u + v)$ which we can calculate efficiently by Euclid's algorithm)

Thus cheating is as hard as prime factorization.

Note that a difficult part is to generate the 100-digit prime. While there are sufficiently many primes to keep trying random numbers until we get one, we need an efficient method to test whether a number is prime. We cannot do this by factorization since it is slow. So we need an efficient prime-checking function.

We can test whether a large number is prime by doing Fermat-like checks. We choose random numbers and take it to the $(p - 1)$th power and see if they become 1. If it is not 1, then it is definitely not a prime. If we do sufficiently many tests that all result in 1, we can be sufficiently certain that it is a prime (even though not with 100% certainty).

(Recent advancements in algorithms have found efficient ways of deterministic prime test, but they are generally slower than the above algorithm and is not widely used)

It is currently believed that it is hard to prime factorize a number, so this is secure as far as we know.

RSA encryption

Theorem (RSA Encryption)**.** We want people to be able to send a message to Bob without Eve eavesdropping. So the message must be encrypted. We want an algorithm that allows anyone to encrypt, but only Bob to decrypt (e.g. many parties sending passwords with the bank).

Let us first agree to write messages as sequences of numbers, e.g. in ASCII or UTF-8.

After encoding, the encryption part is often done with RSA encryption (Rivest, Shamier, Adleman). Bob thinks of two large primes p, q. Let $n = pq$ and pick e coprime to $\phi(n) = (p-1)(q-1)$. Then work out d with $de \equiv 1 \pmod{\phi(n)}$ (i.e. $de = k\phi(n) + 1$). Bob then publishes the pair (n, e).

For Alice to encrypt a message, Alice splits the message into numbers $M < n$. Alice sends $M^e \pmod{n}$ to Bob.

Bob then computes $(M^e)^d = M^{k\phi(n)+1} \equiv M \pmod{n}$ by Fermat-Euler theorem.

How can Eve find M? We can, of course, factorize n, find e efficiently, and be in the same position as Bob. However, it is currently assumed that this is hard. Is there any other way? Currently we do not know if RSA can be broken without factorizing (cf. RSA problem).

6 Real numbers

So far, we have only worked with natural numbers and integers. Unfortunately the real world often involves rational numbers and even real numbers. In this chapter, our goal is to study real numbers. To do so, we will first have to define the real numbers. To do so, we will start from the natural numbers.

Before we start, an important (philosophical) point has to be made. The idea is to define, say, the "real numbers" as a set with some operations (e.g. addition, multiplication) that satisfies some particular properties, known as the *axioms*. When we do this, there are two questions we can ask — is there actually a set that satisfies these properties, and is it unique?

The first question can be answered by an explicit construction, i.e. we find a concrete set that actually satisfies these properties. However, it is important to note that we perform the construction only to show that it makes sense to talk about such a set. For example, we will construct a real number as a pair of subsets of $\mathbb{Q}$, but it would be absurd to actually think that a real number "is" a pair of sets. It's just that it *can be constructed* as a pair of sets. You would be considered insane if you asked if

$$\exists x : x \in 3 \vee x \in \pi$$

holds, even though it is a valid thing to ask if we view each real number as a set (and is in fact true).

The next problem of uniqueness is more subtle. Firstly, it is clear that the constructions themselves aren't unique — instead of constructing the natural number 0 as the set $\emptyset$, as we will later do, we could as well define it as $\{\{\emptyset\}\}$, and the whole construction will go through. However, we could still hope that all possible constructions are "isomorphic" in some way. It turns out this is true for what we have below, but the proofs are not trivial.

However, while this is nice, *it doesn't really matter*. This is since we don't care how, say, the real numbers are constructed. When working with them, we just assume that they satisfy the relevant defining properties. So we can just choose *anything* that satisfies the axioms, and use it. The fact that there are other ones not isomorphic to this is not a problem.

Since we are mostly interested in the natural numbers and the real numbers, we will provide *both* the axiomatic description and explicit construction for these. However, we will only give explicit constructions for the integers and rationals, and we will not give detailed proofs of why they work.

6.1 Construction of numbers

Construction of natural numbers

Our construction of natural numbers will include 0.

Definition (Natural numbers)**.** The natural numbers $\mathbb{N}$ is defined by *Peano's axioms*. We call a set $\mathbb{N}$ "natural numbers" if it has a special element 0 and a map $S : \mathbb{N} \to \mathbb{N}$ that maps n to its "successor" (intuitively, it is $+1$) such that:

(i) $S(n) \neq 0$ for all $n \in \mathbb{N}$

(ii) For all $n, m \in \mathbb{N}$, if $S(n) = S(m)$, then $n = m$.

(iii) For any subset $A \subseteq \mathbb{N}$, if $0 \in A$ and "$n \in A \Rightarrow S(n) \in A$, then in fact $A = \mathbb{N}$.

The last axiom is the axiom of induction.

We write $1 = S(0), 2 = S(1)$, $3 = S(2)$ etc. We can (but will not) prove that the axiom of induction allows us to define functions on $\mathbb{N}$ recursively (cf. IID Logic and Set Theory). Assuming this, we can define addition and multiplication recursively by

$$\begin{aligned} n + 0 &= 0 & n \times 0 &= 0 \\ n + S(m) &= S(n + m) & n \times S(m) &= n \times m + n \end{aligned}$$

We can show by induction that these satisfy the usual rules (e.g. associativity, distributivity).

We can construct this explicitly by $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$ etc. In general, we define $S(n) = \{n\} \cup n$. Note that this is in some sense a circular definition, since we are defining the natural numbers recursively, but to do recursion, we need the natural numbers. Also, it is not clear how we can show this satisfies the axioms above. To actually do this properly, we will need to approach this in a slightly different way, and the details are better left for the IID Logic and Set Theory course.

Construction of integers

Definition (Integers)**.** $\mathbb{Z}$ is obtained from $\mathbb{N}$ by allowing subtraction. Formally, we define $\mathbb{Z}$ to be the equivalence classes of $\mathbb{N} \times \mathbb{N}$ under the equivalence relation

$$(a, b) \sim (c, d) \quad \text{if and only if} \quad a + d = b + c.$$

Intuitively, we think of (a, b) as $a - b$.

We write a for $[(a, 0)]$ and $-a$ for $[(0, a)]$, and define the operations by

$$\begin{aligned} (a, b) + (c, d) &= (a + c, b + d) \\ (a, b) \times (c, d) &= (ac + bd, bd + ad). \end{aligned}$$

We can check that these are well-defined and satisfy the usual properties.

Construction of rationals

Definition (Rationals)**.** $\mathbb{Q}$ is obtained from $\mathbb{Z}$ by allowing division. Formally, we define $\mathbb{Q}$ to be the equivalence classes of $\mathbb{Z} \times \mathbb{N}$ under the relation

$$(a, b) \sim (c, d) \quad \text{if and only if} \quad ad = bc.$$

We write $\frac{a}{b}$ for $[(a, b)]$. We define

$$\begin{aligned} (a, b) + (c, d) &= (ad + bc, bd) \\ (a, b) \times (c, d) &= (ac, bd). \end{aligned}$$

We can check that these are well-defined and satisfy the usual properties.

Algebraically, we say $\mathbb{Q}$ is a "totally ordered field".

Definition (Totally ordered field)**.** A set F equipped with binary operations $+, \times$ and relation $\leq$ is a *totally ordered field* if

(i) F is an additive abelian group with identity 0.

(ii) $F \setminus \{0\}$ is a multiplicative abelian group with identity 1.

(iii) Multiplication is distributed over addition: $a(b + c) = ab + ac$.

(iv) $\leq$ is a total order.

(v) For any $p, q, r \in F$, if $p \leq q$, then $p + r \leq q + r$.

(vi) For any $p, q, r \in F$, if $p \leq q$ and $0 \leq r$, then $pr \leq qr$.

Proposition. $\mathbb{Q}$ is a totally ordered-field.

Examples of non-totally-ordered fields include $\mathbb{Z}_p$, which is a field but not totally ordered.

Proposition. $\mathbb{Q}$ is densely ordered, i.e. for any $p, q \in \mathbb{Q}$, if $p < q$, then there is some $r \in \mathbb{Q}$ such that $p < r < q$.

Proof. Take $r = \frac{p+q}{2}$. □

However, $\mathbb{Q}$ is not enough for our purposes.

Proposition. There is no rational $q \in \mathbb{Q}$ with $q^2 = 2$.

Proof. Suppose not, and $(\frac{a}{b})^2 = 2$, where b is chosen as small as possible. We will derive a contradiction in four ways.

(i) $a^2 = 2b^2$. So a is even. Let $a = 2a'$. Then $b^2 = 2a'^2$. Then b is even as well, and $b = 2b'$. But then $\frac{a}{b} = \frac{a'}{b'}$ with a smaller b'. Contradiction.

(ii) We know that b is a product of primes if $b \neq 1$. Let $p \mid b$. Then $a^2 = 2b^2$. So $p \mid a^2$. So $p \mid a$. Contradict b minimal.

(iii) (Dirichlet) We have $\frac{a}{b} = \frac{2b}{a}$. So $a^2 = 2b^2$. For any, u, v, we have $a^2v = 2b^2v$ and thus $uab + a^2v = uab + 2b^2v$. So $\frac{a}{b} = \frac{au+2bv}{bu+av}$. Put $u = -1, v = 1$. Then $\frac{a}{b} = \frac{2b-a}{a-b}$. Since $a < 2b, a - b < b$. So we have found a rational with smaller b.

(iv) Same as 3, but pick u, v so $bu + av = 1$ since a and b are coprime. So $\frac{a}{b}$ is an integer. □

Construction of real numbers

As shown above, the rational numbers don't have all the numbers we need. What exactly do we mean by "numbers are missing"? We might want to say that the problem is that not all polynomial equations have solutions. However, this is not the real problem. First of all, even if we are working with the reals, not all equations have solutions, e.g. $x^2 + 1 = 0$. Also, some real numbers such as π are not solutions to polynomial equations (with integer coefficients), but we still want them.

The real problem is expressed in terms of least upper bounds, or suprema.

Definition (Least upper bound/supremum and greatest lower bound/infimum)**.** For an ordered set X, $s \in X$ is a *least upper bound* (or *supremum*) for the set $S \subseteq X$, denoted by $s = \sup S$, if

(i) s is an upper bound for S, i.e. for every $x \in S$, we have $x \leq s$.

(ii) if t is any upper bound for S, then $s \leq t$.

Similarly, $s \in X$ is a *greatest lower bound* (or *infimum*) if s is a lower bound and any lower bound $t \leq s$.

By definition, the least upper bound for S, if exists, is unique.

The problem with $\mathbb{Q}$ is that if we let $S = \{q \in \mathbb{Q} : q^2 < 2\}$, then it has no supremum in $\mathbb{Q}$.

Recall that $\mathbb{Q}$ is a totally ordered field. We will define the real numbers axiomatically to be a totally ordered field without this problem.

Definition (Real numbers)**.** The *real numbers* is a totally ordered field containing $\mathbb{Q}$ that satisfies the least upper bound axiom.

Axiom (Least upper bound axiom)**.** Every non-empty set of the real numbers that has an upper bound has a least upper bound.

We have the requirement "non-empty" since every number is an upper bound of $\emptyset$ but it has no least upper bound.

We will leave the construction to the end of the section.

Note that $\mathbb{Q}$ is a subset of $\mathbb{R}$, in the sense that we can find a copy of $\mathbb{Q}$ inside $\mathbb{R}$. By definition of a field, there is a multiplicative identity $1 \in \mathbb{R}$. We can then define the natural numbers by

$$n = \underbrace{1 + \cdots + 1}_{n \text{ times}}.$$

We can then define the negative integers by letting $-n$ be the additive inverse of n. Then $\frac{1}{n}$ is the multiplicative inverse of n (for $n \neq 0$), and $\frac{m}{n}$ is just m copies of $\frac{1}{n}$ added together. This is our canonical copy of $\mathbb{Q}$.

Corollary. Every non-empty set of the real numbers bounded below has an infimum.

Proof. Let S be non-empty and bounded below. Then $-S = \{-x : x \in S\}$ is a non-empty set bounded above, and $\inf S = -\sup(-S)$. □

Alternatively, we can prove it just using the ordering of $\mathbb{R}$:

Proof. Let S be non-empty and bounded below. Let L be the set of all lower bounds of S. Since S is bounded below, L is non-empty. Also, L is bounded above by any member of S. So L has a least upper bound $\sup L$.

For each $x \in S$, we know x is an upper bound of L. So we have $\sup L \leq x$ by definition. So $\sup L$ is indeed a lower bound of S. Also, by definition, every lower bound of S is less than (or equal to) $\sup L$. So this is the infimum. □

Now the set $\{q \in \mathbb{Q} : q^2 < 2\}$ has a supremum in $\mathbb{R}$ (by definition).

We make some useful definitions.

Definition (Closed and open intervals)**.** A *closed interval* $[a, b]$ with $a \leq b \in \mathbb{R}$ is the set $\{x \in \mathbb{R} : a \leq x \leq b\}$.

An *open interval* (a, b) with $a \leq b \in \mathbb{R}$ is the set $\{x \in \mathbb{R} : a < x < b\}$.

Similarly, we can have $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$ and $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$.

Example. Let $S = [0, 1]$. Then $S \neq \emptyset$. Also S has an upper bound, e.g. 2. Hence $\sup S$ exists.

To find it explicitly, notice that 1 is an upper bound for S by definition, and if $t < 1$, then t is not an upper bound for S since $1 \in S$ but $1 \not\leq t$. So every upper bound is at least 1 and therefore 1 is the supremum of S.

Now let $T = (0, 1)$. Again T is non-empty and has an upper bound (e.g. 2). So again $\sup T$ exists. We know that 1 is an upper bound. If $t < 0$, then $0.5 \in S$ but $s \not\leq t$. So t is not an upper bound. Now suppose $0 \leq t < 1$, then $0 < t < \frac{1+t}{2} < 1$ and so $\frac{1+t}{2} \in S$ but $\frac{1+t}{2} \not\leq t$. So t is not an upper bound. So $\sup T = 1$.

Note that these cases differ by $\sup S \in S$ but $\sup T \notin T$. S has a maximum element 1 and the maximum is the supremum. T doesn't have a maximum, but the supremum can still exist.

The real numbers has a rather interesting property.

Theorem (Axiom of Archimedes)**.** Given $r \in \mathbb{R}$, there exists $n \in \mathbb{N}$ with $n > r$.

This was considered an axiom by Archimedes but we can prove this with the least upper bound axiom.

Proof. Assume the contrary. Then r is an upper bound for $\mathbb{N}$. $\mathbb{N}$ is not empty since $1 \in \mathbb{N}$. By the least upper bound axiom, $s = \sup \mathbb{N}$ exists. Since s is the least upper bound for $\mathbb{N}$, $s - 1$ is not an upper bound for $\mathbb{N}$. So $\exists m \in \mathbb{N}$ with $m > s - 1$. Then $m + 1 \in \mathbb{N}$ but $m + 1 > s$, which contradicts the statement that s is an upper bound. □

Notice that every non-empty set $S \in \mathbb{R}$ which is bounded below has a *greatest lower bound* (or *infimum*). In particular, we have

Proposition. $\inf\{\frac{1}{n} : n \in \mathbb{N}\} = 0$.

Proof. Certainly 0 is a lower bound for S. If $t > 0$, there exists $n \in \mathbb{N}$ such that $n \geq 1/t$. So $t \geq 1/n \in S$. So t is not a lower bound for S. □

Theorem. $\mathbb{Q}$ is dense in $\mathbb{R}$, i.e. given $r, s \in \mathbb{R}$, with $r < s$, $\exists q \in \mathbb{Q}$ with $r < q < s$.

Proof. wlog assume first $r \geq 0$ (just multiply everything by -1 if $r < 0$ and swap r and s). Since $s - r > 0$, there is some $n \in \mathbb{N}$ such that $\frac{1}{n} < s - r$. By the Axiom of Archimedes, $\exists N \in \mathbb{N}$ such that $N > sn$.

Let $T = \{k \in \mathbb{N} : \frac{k}{n} \geq s\}$. T is not empty, since $N \in T$. Then by the well-ordering principle, T has a minimum element m. Now $m \neq 1$ since $\frac{1}{n} < s - r \leq s$. Let $q = \frac{m-1}{n}$. Since $m - 1 \notin T$, $q < s$. If $q = \frac{m-1}{n} < r$, then $\frac{m}{n} < r + \frac{1}{n} < s$, so $m \notin T$, contradiction. So $r < q < s$. □

Theorem. There exists $x \in \mathbb{R}$ with $x^2 = 2$.

Proof. Let $S = \{r \in \mathbb{R} : r^2 \leq 2\}$. Then $0 \in S$ so $S \neq \emptyset$. Also for every $r \in S$, we have $r \leq 3$. So S is bounded above. So $x = \sup S$ exists and $0 \leq x \leq 3$.

By trichotomy, either $x^2 < 2, x^2 > 2$ or $x^2 = 2$.

Suppose $x^2 < 2$. Let $0 < t < 1$. Then consider $(x+t)^2 = x^2 + 2xt + t^2 < x^2 + 6t + t \leq x^2 + 7t$. Pick $t < \frac{2-x^2}{7}$, then $(x+t)^2 < 2$. So $x + t \in S$. This contradicts the fact that x is an upper bound of S.

Now suppose $x^2 > 2$. Let $0 < t < 1$. Then consider $(x-t)^2 = x^2 - 2xt + t^2 \geq x^2 - 6t$. Pick $t < \frac{x^2-2}{6}$. Then $(x-t)^2 > 2$, so $x - t$ is an upper bound for S. This contradicts the fact that x is the least upper bound of S.

So by trichotomy, $x^2 = 2$. □

Now let's try to construct the real numbers from the rationals. The idea is that each set like $\{q \in \mathbb{Q} : q^2 < 2\}$ represents a "missing number", namely the supremum $\sqrt{2}$. However, many different sets can correspond to the same missing number, e.g. $\{q \in \mathbb{Q} : q^2 < 2\} \cup \{-3\}$ also "should have" supremum $\sqrt{2}$. So we need to pick a particular one that represents this missing number.

To do so, we pick the maximal one, i.e. a subset S such that for $x \in S$ and $y \in \mathbb{Q}$, if $y < x$, then $y \in S$ (if y is not in S, we can add it to S, and $S \cup \{y\}$ will "have the same upper bound"). Each rational $q \in \mathbb{Q}$ can also be represented by the set $\{x \in \mathbb{Q} : x \leq q\}$. These are known as *Dedekind cuts*. By convention, a Dedekind cut is instead defined as the pair $(S, \mathbb{Q} \setminus S)$, and can be characterized by the following definition:

Definition (Dedekind cut)**.** A *Dedekind cut* of $\mathbb{Q}$ is a set of partition of $\mathbb{Q}$ into L and R such that

$$(\forall l \in L)(\forall r \in R)\, l < r,$$

and R has no minimum.

The requirement that R has no minimum corresponds to our (arbitrary) decision that the rationals should be embedded as

$$q \mapsto \{x \in q : x \leq \mathbb{Q}\}, \{x \in \mathbb{Q} : x > q\},$$

instead of $q \mapsto \{x \in q : x < \mathbb{Q}\}, \{x \in \mathbb{Q} : x \geq q\}$,

We can then construct the set $\mathbb{R}$ from $\mathbb{Q}$ by letting $\mathbb{R}$ be the set of all Dedekind cuts. The supremum of any bounded set of real numbers is obtained by taking the union of (the left sides) of the Dedekind cuts. The definition of the arithmetic operations is left as an exercise for the reader (to actually define them is tedious but not hard).

6.2 Sequences

Here we will look at sequences, with series in the next chapter. Only a brief introduction to these topics will be provided here, as these topics will be studied later in Analysis I.

Definition (Sequence)**.** A *sequence* is a function $\mathbb{N} \to \mathbb{R}$. If a is a sequence, instead of $a(1), a(2), \cdots$, we usually write $a_1, a_2, \cdots$. To emphasize it is a sequence, we write the sequence as (a_n).

We want to capture the notion of a sequence tending to a limit. For example, we want to say that $1, \frac{1}{2}, \frac{1}{3}, \cdots$ tends to 0, while $1, 2, 3, \cdots$ does not converge to a limit.

The idea is that if $a_n \to l$, then we can get as close to l as we like, as long as we are sufficiently far down the sequence. More precisely, given any "error threshold" ε, we can find a (possibly large) number N such that whenever $n \geq N$, we have $|a_n - l| < \varepsilon$.

Definition (Limit of sequence)**.** The sequence (a_n) *tends to* $l \in \mathbb{R}$ as n tends to infinity if and only if

$$(\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall n \geq N)\, |a_n - l| < \varepsilon.$$

If a_n tends to l as n tends to infinity, we write $a_n \to l$ as $n \to \infty$; $\lim_{n\to\infty} a_n = l$; or a_n converges to l.

Intuitively, if $a_n \to l$, we mean given any ε, for sufficiently large n, a_n is always within $l \pm \varepsilon$.

The definition $a_n \not\to l$ is the negation of the above statement:

$$(\exists \varepsilon > 0)(\forall N \in \mathbb{N})(\exists n \geq N)\, |a_n - l| \geq \varepsilon.$$

Definition (Convergence of sequence)**.** The sequence (a_n) *converges* if there exists an l such that $a_n \to l$. The sequence *diverges* if it doesn't converge.

Every proof of $a_n \to l$ looks like: Given $\varepsilon > 0$, (argument to show N exists, maybe depending on ε), such that $\forall n \geq N$, $|a_n - l| < \varepsilon$.

Example. Show that $a_n = 1 - \frac{1}{n} \to 1$.

Given $\varepsilon > 0$, choose $N > \frac{1}{\varepsilon}$, which exists by the Axiom of Archimedes. If $n \geq N$, then $|a_n - 1| = \frac{1}{n} \leq \varepsilon$. So $a_n \to 1$.

Example. Let

$$a_n = \begin{cases} \frac{1}{n} & n \text{ is prime} \\ \frac{1}{2n} & n \text{ is not prime} \end{cases}.$$

We will show that $a_n \to 0$. Given $\varepsilon > 0$. Choose $N > \frac{1}{\varepsilon}$. Then $\forall n \geq N$, $|a_n - 0| \leq \frac{1}{n} < \varepsilon$.

Example. Prove that

$$a_n = \begin{cases} 1 & n \text{ is prime} \\ 0 & n \text{ is not prime} \end{cases}$$

diverges.

Let $\varepsilon = \frac{1}{3}$. Suppose $l \in \mathbb{R}$. If $l < \frac{1}{2}$, then $|a_n - l| > \varepsilon$ when n is prime. If $l \geq \frac{1}{2}$, then $|a_n - l| > \varepsilon$ when n is not prime. Since the primes and non-primes are unbounded, $(\forall N)\exists n > N$ such that $|a_n - l| > \varepsilon$. So a_n diverges.

An important property of $\mathbb{R}$ is the following:

Theorem. Every bounded monotonic sequence converges.

In case of confusion, the terms are defined as follows: (a_n) is increasing if $m \le n$ implies $a_m \le a_n$. Decreasing is defined similarly. Then it is monotonic if it is increasing or decreasing. (a_n) is bounded if there is some $B \in \mathbb{R}$ such that $|a_n| \le B$ for all n.

Proof. wlog assume (a_n) is increasing. The set $\{a_n : n \ge 1\}$ is bounded and non-empty. So it has a supremum l (least upper bound axiom). Show that l is the limit:

Given any $\varepsilon > 0$, $l - \varepsilon$ is not an upper bound of a_n. So $\exists N$ such that $a_N \ge l - \varepsilon$. Since a_n is increasing, we know that $l \ge a_m \ge a_N > l - \varepsilon$ for all $m \ge N$. So $\exists N$ such that $\forall n \ge N$, $|a_n - l| < \varepsilon$. So $a_n \to l$. □

We can show that this theorem is equivalent to the least upper bound axiom.

Definition (Subsequence)**.** A *subsequence* of (a_n) is $(a_{g(n)})$ where $g : \mathbb{N} \to \mathbb{N}$ is strictly increasing. e.g. $a_2, a_3, a_5, a_7 \cdots$ is a subsequence of (a_n).

Theorem. Every sequence has a monotonic subsequence.

Proof. Call a point a_k a "peak" if $(\forall m \ge k)\, a_m \le a_k$. If there are infinitely many peaks, then they form a decreasing subsequence. If there are only finitely many peaks, $\exists N$ such that no a_n with $n > N$ is a peak. Pick a_{N_1} with $N_1 > N$. Then pick a_{N_2} with $N_2 > N_1$ and $a_{N_2} > a_{N_1}$. This is possible because a_{N_1} is not a peak. The pick a_{N_3} with $N_3 > N_2$ and $a_{N_3} > a_{N_2}$, *ad infinitum.* Then we have a monotonic subsequence. □

We will now prove the following basic properties of convergence:

Theorem.

(i) If $a_n \to a$ and $a_n \to b$, then $a = b$ (i.e. limits are unique)

(ii) If $a_n \to a$ and $b_n = a_n$ for all but finitely many n, then $b_n \to a$.

(iii) If $a_n = a$ for all n, then $a_n \to a$.

(iv) If $a_n \to a$ and $b_n \to b$, then $a_n + b_n \to a + b$

(v) If $a_n \to a$ and $b_n \to b$, then $a_n b_n \to ab$

(vi) If $a_n \to a \ne 0$, and $\forall n(a_n \ne 0)$. Then $1/a_n \to 1/a$.

(vii) If $a_n \to a$ and $b_n \to a$, and $\forall n(a_n \le c_n \le b_n)$, then $c_n \to a$. (Sandwich theorem)

Many students are confused as to why we should prove these "obvious" properties of convergence. It seems "obvious" that if a_n converges to a and b_n converges to b, then the sum converges to $a+b$. However, it is not obvious (at least for the first-time learners) that if $(\forall \varepsilon > 0)(\exists N)(\forall n \ge N)\, |a_n - a| < \varepsilon$ and $(\forall \varepsilon > 0)(\exists N)(\forall n \ge N)\, |b_n - b| < \varepsilon$, then $(\forall \varepsilon > 0)(\exists N)(\forall n \ge N)\, |(a_n + b_n) - (a + b)| < \varepsilon$. In some sense, what we are trying to prove is that our attempt at defining convergence actually satisfies the "obvious" properties we think convergence should satisfy.

In proving this, we will make frequent use of the *triangle inequality*: $|x + y| \le |x| + |y|$.

Proof.

(i) Suppose instead $a < b$. Then choose $\varepsilon = \frac{b-a}{2}$. By the definition of the limit, $\exists N_1$ such that $\forall n \geq N_1$, $|a_n - a| < \varepsilon$. There also $\exists N_2$ st. $\forall n \geq N_2$, $|a_n - b| < \varepsilon$.

Let $N = \max\{N_1, N_2\}$. If $n \geq \max\{N_1, N_2\}$, then $|a - b| \leq |a - a_n| + |a_n - b| < 2\varepsilon = b - a$. Contradiction. So $a = b$.

(ii) Given $\varepsilon > 0$, there $\exists N_1$ st. $\forall n \geq N_1$, we have $|a_n - a| < \varepsilon$. Since $b_n = a_n$ for all but finitely many n, there exists N_2 such that $\forall n \geq N_2$, $a_n = b_n$.

Let $N = \max\{N_1, N_2\}$. Then $\forall n \geq N$, we have $|b_n - a| = |a_n - a| < \varepsilon$. So $b_n \to a$.

(iii) $\forall \varepsilon$, take $N = 1$. Then $|a_n - a| = 0 < \varepsilon$ for all $n \geq 1$.

(iv) Given $\varepsilon > 0$, $\exists N_1$ such that $\forall n \geq N_1$, we have $|a_n - a| < \varepsilon/2$. Similarly, $\exists N_2$ such that $\forall n \geq N_2$, we have $|b_n - b| < \varepsilon/2$.

Let $N = \max\{N_1, N_2\}$. Then $\forall n \geq N$, $|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \varepsilon$.

(v) Given $\varepsilon > 0$, Then there exists N_1, N_2, N_3 such that

$$\begin{aligned}
&\forall n \geq N_1 : |a_n - a| < \frac{\varepsilon}{2(|b| + 1)}\\
&\forall n \geq N_2 : |b_n - b| < \frac{\varepsilon}{2|a|}\\
&\forall n \geq N_3 : |b_n - b| < 1 \Rightarrow |b_n| < |b| + 1
\end{aligned}$$

Then let $N = \max\{N_1, N_2, N_3\}$. Then $\forall n \geq N$,

$$\begin{aligned}
|a_n b_n - ab| &= |b_n(a_n - a) + a(b_n - b)|\\
&\leq |b_n||a_n - a| + |a||b_n - b|\\
&< (|b| + 1)|a_n - a| + |a||b_n - b|\\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}\\
&= \varepsilon
\end{aligned}$$

(vi) Given $\varepsilon > 0$, then $\exists N_1, N_2$ such that $|a_n - a| < \frac{|a|^2}{2}\varepsilon$ and $|a_n - a| < \frac{|a|}{2}$.
Let $N = \max\{N_1, N_2\}$. The $\forall n \geq N$,

$$\begin{aligned}
\left|\frac{1}{a_n} - \frac{1}{a}\right| &= \frac{|a_n - a|}{|a_n||a|}\\
&< \frac{2}{|a|^2}|a_n - a|\\
&< \varepsilon
\end{aligned}$$

(vii) By (iii) to (v), we know that $b_n - a_n \to 0$. Let $\varepsilon > 0$. Then $\exists N$ such that $\forall n \geq N$, we have $|b_n - a_n| < \varepsilon$. So $|c_n - a_n| < \varepsilon$. So $c_n - a_n \to 0$. So $c_n = (c_n - a_n) + a_n \to a$. □

Example. Let $x_n = \frac{n^2(n+1)(2n+1)}{n^4+1}$. Then we have

$$x_n = \frac{(1+1/n)(2+1/n)}{1+1/n^4} \to \frac{1\cdot 2}{1} = 2$$

by the theorem (many times).

Example. Let $y_n = \frac{100^n}{n!}$. Since $\frac{y_{n+1}}{y_n} = \frac{100}{n+1} < \frac{1}{2}$ for large $n > 200$, we know that $0 \le y_n < y_{200} \cdot \frac{2^{200}}{2^n}$. Since $y_{200} \cdot \frac{2^{200}}{2^n} \to 0$, we know that $y_n \to 0$ as well.

6.3 Series

In a field, the sum of two numbers is defined. By induction, the sum of finitely many numbers is defined as well. However, infinite sums ("series") are not. We will define what it means to take an infinite sum. Of course, infinite sums exist only for certain nice sums. For example, $1 + 1 + 1 + \cdots$ does not exist.

Definition (Series and partial sums)**.** Let (a_n) be a sequence. Then $s_m = \sum_{n=1}^m a_n$ is the *mth partial sum* of (a_n). We write

$$\sum_{n=1}^\infty a_n = \lim_{m\to\infty} s_m$$

if the limit exists.

Example. Let $a_n = \frac{1}{n(n-1)}$ for $n \ge 2$. Then

$$s_m = \sum_{n=2}^m \frac{1}{n(n-1)} = \sum_{n=2}^m \left(\frac{1}{n-1} - \frac{1}{n}\right) = 1 - \frac{1}{m} \to 1.$$

Then

$$\sum_{n=2}^\infty \frac{1}{n(n-1)} = 1.$$

Example. Let $a_n = \frac{1}{n^2}$. Then $s_m = \sum_{n=1}^m \frac{1}{n^2}$. We know that s_m is increasing. We also know that $s_m \le 1 + \sum \frac{1}{n(n-1)} \le 2$, i.e. it is bounded above. So s_m converges and $\sum_{n=1}^\infty \frac{1}{n^2}$ exists (in fact it is $\pi^2/6$).

Example. (Geometric series) Suppose $a_n = r^n$, where $|r| < 1$. Then $s_m = r \cdot \frac{1-r^m}{1-r} \to \frac{r}{1-r}$ since $r^n \to 0$. So

$$\sum_{n=1}^\infty r^n = \frac{r}{1-r}.$$

Example. (Harmonic series) Let $a_n = \frac{1}{n}$. Consider

$$\begin{aligned} S_{2^k} &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \cdots + \frac{1}{2^k} \\ &\ge 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{16} + \cdots + \frac{1}{2^k} \\ &\ge 1 + \frac{k}{2}. \end{aligned}$$

So $\sum_{n=1}^\infty \frac{1}{n}$ diverges.

Decimal expansions

Definition (Decimal expansion)**.** Let (d_n) be a sequence with $d_n \in \{0, 1, \cdots 9\}$. Then $\sum_{n=1}^{\infty} \frac{d}{10^n}$ converges to a limit r with $0 \le r \le 1$ since the partial sums s_m are increasing and bounded by $\sum \frac{9}{10^n} \to 1$ (geometric series). We say $r = 0.d_1d_2d_3\cdots$, the *decimal expansion* of r.

Does every x with $0 \le x < 1$ have a decimal expansion? Pick d_1 maximal such that $\frac{d_1}{10} \le x < 1$. Then $0 \le x - \frac{d_1}{10} < \frac{1}{10}$ since d_1 is maximal. Then pick d_2 maximal such that $\frac{d_2}{100} \le x - \frac{d_1}{10}$. By maximality, $0 \le x - \frac{d_1}{10} - \frac{d_2}{100} < \frac{1}{100}$. Repeat inductively, pick maximal d_n with

$$\frac{d_n}{10^n} \le x - \sum_{j=1}^{n-1} \frac{d_j}{10^j}$$

so

$$0 \le x - \sum_{j=1}^{n} \frac{d_j}{10^j} < \frac{1}{10^n}.$$

Since both LHS and RHS $\to 0$, by sandwich, $x - \sum_{j=1}^{\infty} \frac{d_j}{10^j} = 0$, i.e. $x = 0.d_1d_2\cdots$.

Since we have shown that at least one decimal expansion, can the same number have two different decimal expansions? i.e. if $0.a_1a_2\cdots = 0.b_1b_2\cdots$, must $a_i = b_i$ for all i?

Now suppose that the a_j and b_j are equal until k, i.e. $a_j = b_j$ for $j < k$. wlog assume $a_k < b_k$. Then

$$\sum_{j=k+1}^{\infty} \frac{a_j}{10^j} \le \sum_{j=k+1}^{\infty} \frac{9}{10^j} = \frac{9}{10^{k+1}} \cdot \frac{1}{1 - 1/10} = \frac{1}{10^k}.$$

So we must have $b_k = a_k + 1$, $a_j = 9$ for $j > k$ and $b_j = 0$ for $j > k$. For example, $0.47999\cdots = 0.48000\cdots$.

6.4 Irrational numbers

Recall $\mathbb{Q} \subseteq \mathbb{R}$.

Definition (Irrational number)**.** Numbers in $\mathbb{R} \setminus \mathbb{Q}$ are *irrational*.

Definition (Periodic number)**.** A decimal is *periodic* if after a finite number ℓ of digits, it repeats in blocks of k for some k, i.e. $d_{n+k} = d_n$ for $n > \ell$.

Proposition. A number is periodic iff it is rational.

Proof. Clearly a periodic decimal is rational: Say $x = 0.7413157157157\cdots$. Then

$$\begin{aligned} 10^\ell x &= 10^4 x \\ &= 7413.157157\cdots \\ &= 7413 + 157\left(\frac{1}{10^3} + \frac{1}{10^6} + \frac{1}{10^9} + \cdots\right) \\ &= 7413 + 157 \cdot \frac{1}{10^3} \cdot \frac{1}{1 - 1/10^3} \in \mathbb{Q} \end{aligned}$$

Conversely, let $x \in \mathbb{Q}$. Then x has a periodic decimal. Suppose $x = \frac{p}{2^c 5^d q}$ with $(q, 10) = 1$. Then $10^{\max(c,d)} x = \frac{a}{q} = n + \frac{b}{q}$ for some $a, b, n \in \mathbb{Z}$ and $0 \leq b < q$. However, since $(q, 10) = 1$, by Fermat-Euler, $10^{\phi(q)} \equiv 1 \pmod{q}$, i.e. $10^{\phi(q)} - 1 = kq$ for some k. Then

$$\frac{b}{q} = \frac{kb}{kq} = \frac{kb}{999\cdots 9} = kb\left(\frac{1}{10^{\phi(q)}} + \frac{1}{10^{2\phi(q)}} + \cdots\right).$$

Since $kb < kq < 10^{\phi(q)}$, write $kb = d_1 d_2 \cdots d_{\phi(q)}$. So $\frac{b}{q} = 0.d_1 d_2 \cdots d_{\phi(q)} d_1 d_2 \cdots$ and x is periodic. □

Example. $x = 0.01101010001010\cdots$, where 1s appear in prime positions, is irrational since the digits don't repeat.

6.5 Euler's number

Definition (Euler's number)**.**

$$e = \sum_{j=0}^{\infty} \frac{1}{j!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots$$

This sum exists because the partial sums are bounded by $1 + \frac{1}{1} + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \cdots = 3$ and it is increasing. So $2 < e < 3$.

Proposition. e is irrational.

Proof. Is $e \in \mathbb{Q}$? Suppose $e = \frac{p}{q}$. We know $q \geq 2$ since e is not an integer (it is between 2 and 3). Then $q!e \in \mathbb{N}$. But

$$q!e = \underbrace{q! + q! + \frac{q!}{2!} + \frac{q!}{3!} + \cdots + \frac{q!}{q!}}_{n} + \underbrace{\frac{q!}{(q+1)!} + \frac{q!}{(q+2)!} + \cdots}_{x},$$

where $n \in \mathbb{N}$. We also have

$$x = \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \cdots.$$

We can bound it by

$$0 < x < \frac{1}{q+1} + \frac{1}{(q+1)^2} + \frac{1}{(q+1)^3} + \cdots = \frac{1}{q+1} \cdot \frac{1}{1 - 1/(q+1)} = \frac{1}{q} < 1.$$

This is a contradiction since $q!e$ must be in $\mathbb{N}$ but it is a sum of an integer n plus a non-integer x. □

6.6 Algebraic numbers

Rational numbers are "nice", because they can be written as fractions. Irrational numbers are bad. However, some irrational numbers are worse than others. We can further classify some irrational numbers as being *transcendental*.

Definition (Algebraic and transcendental numbers)**.** An *algebraic number* is a root of a polynomial with integer coefficients (or rational coefficients). A number is *transcendental* if it is not algebraic.

Proposition. All rational numbers are algebraic.

Proof. Let $x = \frac{p}{q}$, then x is a root of $qx - p = 0$. □

Example. $\sqrt{2}$ is irrational but algebraic since it is a root of $x^2 - 2 = 0$.

So do transcendental numbers exist?

Theorem. (Liouville 1851; Non-examinable) L is transcendental, where

$$L = \sum_{n=1}^{\infty} \frac{1}{10^{n!}} = 0.11000100\cdots$$

with 1s in the factorial positions.

Proof. Suppose instead that $f(L) = 0$ where $f(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_0$, where $a_i \in \mathbb{Z}$, $a_k \neq 0$.

For any rational p/q, we have

$$f\left(\frac{p}{q}\right) = a_k\left(\frac{p}{q}\right)^k + \cdots + a_0 = \frac{\text{integer}}{q^k}.$$

So if p/q is not a root of f, then $|f(p/q)| \geq q^{-k}$.

For any m, we can write $L =$ first m terms + rest of the terms $= s + t$.

Now consider $|f(s)| = |f(L) - f(s)|$ (since $f(L) = 0$). We have

$$\begin{aligned}
|f(L) - f(s)| &= \left|\sum a_i(L^i - s^i)\right| \\
&\leq \sum |a_i(L^i - s^i)| \\
&= \sum |a_i|(L - s)(L^{i-1} + \cdots + s^{i-1}) \\
&\leq \sum |a_i|(L - s)i, \\
&= (L - s)\sum i|a_i| \\
&= tC
\end{aligned}$$

with $C = \sum i|a_i|$.

Writing s as a fraction, its denominator is at most $10^{m!}$. So $|f(s)| \geq 10^{-k\times m!}$. Combining with the above, we have $tC \geq 10^{-k\times m!}$.

We can bound t by

$$t = \sum_{j=m+1}^{\infty} 10^{-j!} \leq \sum_{\ell=(m+1)!}^{\infty} 10^{-\ell} = \frac{10}{9}10^{-(m+1)!}.$$

So $(10C/9)10^{-(m+1)!} \geq 10^{-k\times m!}$. Pick $m \in \mathbb{N}$ so that $m > k$ and $10^{m!} > \frac{10C}{9}$. This is always possible since both k and $10C/9$ are constants. Then the inequality gives $10^{-(m+1)} \geq 10^{-(k+1)}$, which is a contradiction since $m > k$. □

Theorem. (Hermite 1873) e is transcendental.

Theorem. (Lindermann 1882) π is transcendental.

7 Countability

After messing with numbers, we finally get back to sets. Here we are concerned about the sizes of sets. We can count how big a set is by constructing bijections. Two sets have the same number of things if there is a bijection between them. In particular, a set has n things if we can bijection it with $[n] = \{1, 2, 3, \cdots, n\}$.

First first prove a few preliminary properties about bijecting with $[n]$ that should be obviously true.

Lemma. If $f : [n] \to [n]$ is injective, then f is bijective.

Proof. Perform induction on n: It is true for $n = 1$. Suppose $n > 1$. Let $j = f(n)$. Define $g : [n] \to [n]$ by

$$g(j) = n, \quad g(n) = j, \quad g(i) = i \text{ otherwise.}$$

Then g is a bijection. So the map $g \circ f$ is injective. It fixes n, i.e. $g \circ f(n) = n$. So the map $h : [n-1] \to [n-1]$ by $h(i) = g \circ f(i)$ is well-defined and injective. So h is surjective. So h is bijective. So $g \circ f$ is bijective. So is f. □

Corollary. If A is a set and $f : A \to [n]$ and $g : A \to [m]$ are both bijections, then $m = n$.

Proof. wlog assume $m \geq n$. Let $h : [n] \to [m]$ with $h(i) = i$, which is injective. Then the map $h \circ f \circ g^{-1} : [m] \to [m]$ is injective. Then by the lemma this is surjective. So h must be surjective. So $n \geq m$. Hence $n = m$. □

This shows that we cannot biject a set to two different numbers, or a set cannot have two different sizes!

Definition (Finite set and cardinality of set)**.** The set A is *finite* if there exists a bijection $A \to [n]$ for some $n \in \mathbb{N}_0$. The *cardinality* or *size* of A, written as $|A|$, is n. By the above corollary, this is well-defined.

Lemma. Let $S \subseteq \mathbb{N}$. Then either S is finite or there is a bijection $g : \mathbb{N} \to S$.

Proof. If $S \neq \emptyset$, by the well-ordering principle, there is a least element $s_1 \in S$. If $S \setminus \{s_1\} \neq \emptyset$, it has a least element s_2. If $S \setminus \{s_1, s_2\}$ is not empty, there is a least element s_3. If at some point the process stops, then $S = \{s_1, s_2, \cdots, s_n\}$, which is finite. Otherwise, if it goes on forever, the map $g : \mathbb{N} \to S$ given by $g(i) = s_i$ is well-defined and is an injection. It is also a surjection because if $k \in S$, then k is a natural number and there are at most k elements of S less than k. So k will be mapped to s_i for some $i \leq k$. □

Definition (Countable set)**.** A set A is *countable* if A is finite or there is a bijection between A and $\mathbb{N}$. A set A is *uncountable* if A is not countable.

This is one possible definition of countability, but there are some (often) more helpful definitions.

Theorem. The following are equivalent:

(i) A is countable

(ii) There is an injection from $A \to \mathbb{N}$

(iii) $A = \emptyset$ or there is a surjection from $\mathbb{N} \to A$

Proof. (i) $\Rightarrow$ (iii): If A is finite, there is a bijection $f : A \to S$ for some $S \subseteq \mathbb{N}$. For all $x \in \mathbb{N}$, if $x \in S$, then map $x \mapsto f^{-1}(x)$. Otherwise, map x to any element of A. This is a surjection since $\forall a \in A$, we have $f(a) \mapsto a$.

(iii) $\Rightarrow$ (ii): If $A \neq \emptyset$ and $f : \mathbb{N} \to A$ is a surjection. Define a map $g : A \to \mathbb{N}$ by $g(a) = \min f^{-1}(\{a\})$, which exists by well-ordering. So g is an injection.

(ii) $\Rightarrow$ (i): If there is an injection $f : A \to \mathbb{N}$, then f gives a bijection between A and $S = f(A) \subseteq \mathbb{N}$. If S is finite, so is A. If S is infinite, there is a bijection g between S and $\mathbb{N}$. So there is a bijection $g \circ f$ between A and $\mathbb{N}$. □

Often, the injection definition is the most helpful.

Proposition. The integers $\mathbb{Z}$ are countable.

Proof. The map $f : \mathbb{Z} \to \mathbb{N}$ given by

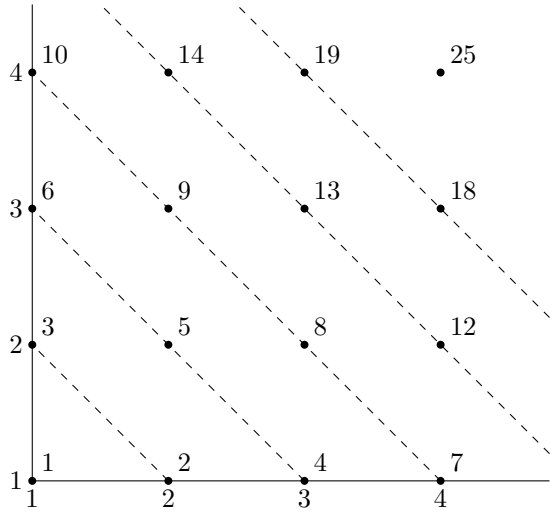
$$f(n) = \begin{cases} 2n & n > 0 \\ 2(-n) + 1 & n \leq 0 \end{cases}$$

is a bijection. □

Proposition. $\mathbb{N} \times \mathbb{N}$ is countable.

Proof. We can map $(a, b) \mapsto 2^a 3^b$ injectively by the fundamental theorem of arithmetic. So $\mathbb{N} \times \mathbb{N}$ is countable.

We can also have a bijection by counting diagonally: $(a, b) \mapsto \binom{a+b}{2} - a + 1$:



□

Since $\mathbb{Z}$ is countable, we have an injection $\mathbb{Z} \to \mathbb{N}$, so there is an injection from $\mathbb{Z} \times \mathbb{N} \to \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. So $\mathbb{Z} \times \mathbb{N}$ is countable. However, the rationals are the equivalence classes of $\mathbb{Z} \times \mathbb{N}$. So $\mathbb{Q}$ is countable.

Proposition. If $A \to B$ is injective and B is countable, then A is countable (since we can inject $B \to \mathbb{N}$).

Proposition. $\mathbb{Z}^k$ is countable for all $k \in \mathbb{N}$

Proof. Proof by induction: $\mathbb{Z}$ is countable. If $\mathbb{Z}^k$ is countable, $\mathbb{Z}^{k+1} = \mathbb{Z} \times \mathbb{Z}^k$. Since we can map $\mathbb{Z}^k \to \mathbb{N}$ injectively by the induction hypothesis, we can map injectively $\mathbb{Z}^{k+1} \to \mathbb{Z} \times \mathbb{N}$, and we can map that to $\mathbb{N}$ injectively. □

Theorem. A countable union of countable sets is countable.

Proof. Let I be a countable index set, and for each $\alpha \in I$, let A_α be a countable set. We need to show that $\bigcup_{\alpha\in I} A_\alpha$ is countable. It is enough to construct an injection $h : \bigcup_{\alpha\in I} A_\alpha \to \mathbb{N} \times \mathbb{N}$ because $\mathbb{N} \times \mathbb{N}$ is countable. We know that I is countable. So there exists an injection $f : I \to \mathbb{N}$. For each $\alpha \in I$, there exists an injection $g_\alpha : A_\alpha \to \mathbb{N}$.

For $a \in \bigcup A_\alpha$, pick $m = \min\{j \in \mathbb{N} : a \in A_\alpha \text{ and } f(\alpha) = j\}$, and let α be the corresponding index such that $f(\alpha) = m$. We then set $h(a) = (m, g_\alpha(a))$, and this is an injection. □

Proposition. $\mathbb{Q}$ is countable.

Proof. It can be proved in two ways:

(i) $\mathbb{Q} = \bigcup_{n\geq 1} \frac{1}{n}\mathbb{Z} = \bigcup_{n\geq 1} \left\{\frac{m}{n} : m \in \mathbb{Z}\right\}$, which is a countable union of countable sets.

(ii) $\mathbb{Q}$ can be mapped injectively to $\mathbb{Z} \times \mathbb{N}$ by $a/b \mapsto (a, b)$, where $b > 0$ and $(a, b) = 1$. □

Theorem. The set of algebraic numbers is countable.

Proof. Let $\mathcal{P}_k$ be the set of polynomials of degree k with integer coefficients. Then $a_k x^k + a_{k-1}x^{k-1} + \cdots + a_0 \mapsto (a_k, a_{k-1}, \cdots, a_0)$ is an injection $\mathcal{P}_k \to \mathbb{Z}^{k+1}$. Since $\mathbb{Z}^{k+1}$ is countable, so is $\mathcal{P}_k$.

Let $\mathcal{P}$ be the set of all polynomials with integer coefficients. Then clearly $\mathcal{P} = \bigcup \mathcal{P}_k$. This is a countable union of countable sets. So $\mathcal{P}$ is countable.

For each polynomial $p \in \mathcal{P}$, let R_p be the set of its roots. Then R_p is finite and thus countable. Hence $\bigcup_{p\in\mathcal{P}} R_p$, the set of all algebraic numbers, is countable. □

Theorem. The set of real numbers $\mathbb{R}$ is uncountable.

Proof. (Cantor's diagonal argument) Assume $\mathbb{R}$ is countable. Then we can list the reals as $r_1, r_2, r_3 \cdots$ so that every real number is in the list. Write each r_n uniquely in decimal form (i.e. without infinite trailing '9's). List them out vertically:

$$\begin{aligned} r_1 &= n_1 \,.\, d_{11}\, d_{12}\, d_{13}\, d_{14} \cdots \\ r_2 &= n_2 \,.\, d_{21}\, d_{22}\, d_{23}\, d_{24} \cdots \\ r_3 &= n_3 \,.\, d_{31}\, d_{32}\, d_{33}\, d_{34} \cdots \\ r_4 &= n_4 \,.\, d_{41}\, d_{42}\, d_{43}\, d_{44} \cdots \end{aligned}$$

Define $r = 0 \,.\, d_1\, d_2\, d_3\, d_4 \cdots$ by $d_n = \begin{cases} 0 & d_{nn} \neq 0 \\ 1 & d_{nn} = 0 \end{cases}$. Then by construction, this differs from the nth number in the list by the nth digit, and is so different from every number in the list. Then r is a real number but not in the list. Contradiction. □

Corollary. There are uncountable many transcendental numbers.

Proof. If not, then the reals, being the union of the transcendentals and algebraic numbers, must be countable. But the reals is uncountable. □

This is an easy but non-constructive proof that transcendental numbers exists. "If we can't find one, find lots!" (it is debatable whether this proof is constructive or not. Some argue that we can use this to construct a transcendental number by listing all the algebraic numbers and perform the diagonal argument to obtain a number not in the list, i.e. a transcendental number. So this is in fact constructive)

Example. Let $\mathcal{F}_k = \{Y \subseteq \mathbb{N} : |Y| = k\}$, i.e. the set of all subsets of $\mathbb{N}$ of size k. We can inject $\mathcal{F}_k \to \mathbb{Z}^k$ in the obvious way, e.g. $\{1, 3, 7\} \mapsto (1, 3, 7)$ etc. So it is countable. So $\mathcal{F} = \bigcup_{k\geq 0} \mathcal{F}_k$, the set of all finite subsets of $\mathbb{N}$ is countable.

Example. Recall $\mathcal{P}(X) = \{Y : Y \subseteq X\}$. Now suppose $\mathcal{P}(\mathbb{N})$ is countable. Let $S_1, S_2, S_3, \cdots$ be the list of all subsets of $\mathbb{N}$. Let $S = \{n : n \notin S_n\}$. But then S is not in the list. Contradiction. So $\mathcal{P}(\mathbb{N})$ is uncountable.

Example. Let Σ be the set of all functions $\mathbb{N} \to \mathbb{N}$ (i.e. the set of all integer sequences). If Σ were countable, we could list it as $f_1, f_2, f_3 \cdots$. But then consider f given by $f(n) = \begin{cases} 1 & f_n(n) \neq 1 \\ 2 & f_n(n) = 1 \end{cases}$. Again f is not in the list. Contradiction. So Σ is uncountable.

Alternatively, there is a bijection between $\mathcal{P}(\mathbb{N})$ and the set of 0, 1 sequences by $S \mapsto$ the indicator function. So we can inject $\mathcal{P}(\mathbb{N}) \to \Sigma$ by $S \mapsto$ indicator function $+1$. So Σ cannot be countable (since $\mathcal{P}(\mathbb{N})$ is uncountable).

Or, we can let $\Sigma^* \subseteq \Sigma$ be the set of bijections from $\mathbb{N} \to \mathbb{N}$. Let $\Sigma^{**} \subseteq \Sigma^*$ be the bijections of the special form: for every n,

$$\text{either } \begin{cases} f(2n-1) = 2n-1 \\ f(2n) = 2n \end{cases} \text{, or } \begin{cases} f(2n-1) = 2n \\ f(2n) = 2n-1 \end{cases},$$

i.e. for every odd-even pair, we either flip them or keep them the same.

But there is a bijection between Σ^{**} and $0, 1$ sequences: if the nth term in the sequence $= 0$, don't flip the nth pair in the function, vice versa. Hence Σ^{**} is uncountable.

Theorem. Let A be a set. Then there is no surjection from $A \to \mathcal{P}(A)$.

Proof. Suppose $f : A \to \mathcal{P}(A)$ is surjective. Let $S = \{a \in A : a \notin f(a)\}$. Since f is surjective, there must exist $s \in A$ such that $f(s) = S$. If $s \in S$, then $s \notin S$ by the definition of S. Conversely, if $s \notin S$, then $s \in S$. Contradiction. So f cannot exist. □

This shows that there are infinitely many different possible "infinite sizes" of sets.

We conclude by two theorems that we will not prove.

Theorem (Cantor-Schröder-Bernstein theorem)**.** Suppose there are injections $A \to B$ and $B \to A$. Then there's a bijection $A \leftrightarrow B$.

Continuum hypothesis. There is no set whose size lies between $\mathbb{N}$ and $\mathbb{R}$. In 1963, Paul Cohen proved that it is impossible to prove this or disprove this statement (in ZFC). The proof can be found in the Part III Topics in Set Theory course.